PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 9/10, 15/82

A2

(11) International Publication Number: **WO 00/11142**

(43) International Publication Date: 2 March 2000 (02.03.00)

(21) International Application Number: PCT/EP99/06037

(22) International Filing Date: 18 August 1999 (18.08.99)

(30) Priority Data:
09/137,434 20 August 1998 (20.08.98) US

(71) Applicant *(for all designated States except US)*: CULTOR CORPORATION [FI/FI]; Keilaranta 9, FIN–02150 Espoo (FI).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: REINIKAINEN, Tapani [FI/FI]; Pajalahdentie 21 A 16, FIN–00200 Helsinki (FI). NYYSSÖLÄ, Antti [FI/FI]; Messeniuksenkatu 8 A 8, FIN–00250 Helsinki (FI). KEROVUO, Janne [FI/FI]; Merimiehenkatu 39 B 36, FIN–00150 Helsinki (FI).

(74) Agents: KOLB, Helga et al.; Hoffmann . Eitle, Arabellastrasse 4, D–81925 Munich (DE).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: METHYLTRANSFERASES, NUCLEIC ACID MOLECULES ENCODING METHYLTRANSFERASES, THEIR RECOMBINANT EXPRESSION AND USES THEREOF

(57) Abstract

The present invention relates to proteins which are capable of functioning as methyltransferases. More, specifically, the present invention relates to methyltransferases which are capable of carrying out at least one of the following reactions: the conversion of glycine to sarcosine, sarcosine to dimethyl glycine and dimethyl glycine to betaine in the presence of a methyl group donor. Furthermore, the present invention relates to nucleic acid molecules encoding such methyltransferase proteins, recombinant organisms which are capable of expressing said nucleic acids as well as the use of said recombinant organisms.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Methyltransferases, nucleic acid molecules encoding methyltransferases, their recombinant expression and uses ther of

The present invention relates to proteins which are capable of functioning as methyltransferases. More, specifically, the present invention relates to methyltransferases which are capable of carrying out at least one of the following reactions: the conversion of glycine to sarcosine (N-methylglycine), sarcosine to N,N-dimethyl glycine and N,N-dimethyl glycine to betaine (N,N,N-trimethylglycine) in the presence of a methyl group donor. Furthermore, the present invention relates to nucleic acid molecules encoding such methyltransferase proteins, recombinant organisms which are capable of expressing said nucleic acids as well as the use of said recombinant organisms.

Technological background

Betaine (N,N,N-trimethylglycine) is a quaternary ammonium compound which can be found in many micro-organisms, animals and plants. Betaine is synthesized or accumulated in living cells in response to abiotic stress (salinity, desiccation or low temperatures) (Mc Gue and Hanson, 1990; Csonka, 1989; Yancey et al., 1982; Wyn Jones et al., 1977; Gorham, 1995; Bohnert and Jensen, 1996). Due to its physical properties, betaine is an osmolyte and thus it is able to restore and maintain osmotic balance of living cells. In addition, it has been demonstrated that betaine stabilizes and protects cell membranes (Coughlan and Heber, 1982) and other macromolecules (i.e. enzymes) in the cell (Papagiorgiou and Murata, 1995).

Betaine is synthesized by a number of microbes. In addition, these microbes are usually capable of accumulating betaine or a precursor, choline, from the culture medium (Boch et al., 1994; Perroud and Rudulier, 1985; Kempf and Bremer, 1995; Glaasker et al., 1996; Peter et al., 1996; and Kappes et al., 1996). Practically all halophilic bacteria are able to use

betaine as at least one of their osmolytes in order to survive in the high ionic strength environment.

Also, many plants synthesize betaine in response to drought or salinity (Rhodes and Hanson, 1993; Gorham, 1995). A correlation between cold tolerance and betaine synthesis has been demonstrated (Holmberg, N. 1996, Kishitani et al., 1994; Nomura et al., 1995). Intracellular concentrations as low as 1 mM have shown to give protective effects in plants (Ishitani et al., 1993). Increasing the betaine content of plants by genetic engineering or plant breeding has shown to result in more salt or cold tolerant transgenic plants. A betaine biosynthesis pathway has been introduced, e.g. in tobacco (Lilius et al. 1996; JP 8-266179; JP 8-103267) Arabidopsis (Hayashi et al., 1997) and rice (Nakamura et al., 1997; Guo, 1997). However, the results obtained are very preliminary and only phenotypic effects (salt or cold tolerance) have been demonstrated. Very little is known with respect to the relationship between the concentration dependence of the intracellular betaine and stress tolerance.

It has recently been shown that plants can acquire better stress tolerance by accumulating exogenously applied betaine. Foliar spraying of betaine in a specific phase of growth has been shown to increase the productivity of many crops. (Mäkelä et al., 1996; Agboma, 1997). Typically, 15 % crop yield improvements have been obtained with many species (sorghum, maize, soybean, cotton, potato, tomato) under conditions of salinity and drought. The detailed physiological mechanism of betaine action is not fully understood, but it is known that betaine stimulates photosynthesis and decreases photorespiration.

In animal cells, betaine also acts as a methyl group donor. The most important function is to the ability to methylate homocysteine back to methionine, which can then further be used as a methyl group donor when metabolized to S-adenosyl methionine (SAM). It has been demonstrated that orally

administered betaine relieves diarrhea and dehydration in many animals and inhibits invasion of gut epithelium by coccidia parasite (Ferket, 1994; Augustine and McNaughton, 1996). As a methyl group donor betaine has been shown to be lipotropic, thus decreasing the amount of fat in chicken meat (Saunderson and MacKinlay, 1990; Barak et al., 1993).

The most extensively studied betaine biosynthesis pathway is the two-step oxidation reaction of choline to betaine via betaine aldehyde. This metabolic pathway has been demonstrated to exist in number of microbes, plants and animal cells. The choline-betaine pathway of E. coli (Lamark et al., 1991) and Pseudomonas aeruginosa (Nagasawa et al., 1976) comprises an oxygen-dependent choline dehydrogenase, which catalyzes the oxidation of both choline to betaine aldehyde and betaine aldehyde to betaine. The E. coli choline dehydrogenase gene has been cloned and sequenced. In addition, it has successfully been expressed in many heterologous organisms (Nomura et al., 1995; Lilius et al., 1996; Hayashi et al., 1997; Nakamura et al., 1997; Guo, 1997). The enzyme is membrane-bound. The reaction is independent of soluble cofactors and electron-transfer linked. Choline dehydrogenase genes from Sinorhizobium meliloti (Pocard et al., 1997) and Bacillus (Boch et al., 1996) have also been isolated.

Alternatively, the oxidation of choline to betaine can be catalyzed by a choline oxidase found for example in some Corynebacteria, Brevibacterium and Alcaligenes species (Nakanishi and Machida, 1981; Kojima et al., 1987). The enzyme has been shown to also exist in some fungal strains (Tani et al., 1979). The choline oxidase of Arthrobacter pascens has been cloned and successfully expressed in E. coli (Rozwadovski et al., 1991) or Synechococcus (Deshnium et al., 1995). The reaction uses molecular oxygen as the hydrogen acceptor and hydrogen peroxide is formed in the reaction.

In plants, the synthesis of betaine pathway has been investigated in detail in sugar beet and spinach (McCue et al., 1992; Weretilnyk and Hanson, 1989). The first step is catalyzed by a choline mono-oxygenase. In plants the enzyme is located in the chloroplast stroma (Brouquisse et al., 1989). The gene has recently been cloned from spinach (Rathinasabapathi et al., 1997).

The second oxidation step from betaine aldehyde to betaine may also be catalyzed by a betaine aldehyde dehydrogenase. A betaine aldehyde dehydrogenase has also been found in a number of organisms (Pseudomonas aeruginosa (30)). Also some plants have this enzyme (Hanson et al., 1985). In plants, the enzyme has been demonstrated to have wider substrate specificity and thus it also catalyzes other reactions (Trossat et al., 1997). The gene has been cloned from E. coli (Lamark et al., 1991), spinach (Weretilnyk and Hanson, 1990) and also from barley (Ishitani et al., 1995). The E. coli gene has also successfully been expressed in transgenic tobacco (Holmström et al., 1994).

Ability to synthesize betaine de novo is rare among aerobic heterotrophic eubacteria. Of all strains examined, only Actinopolyspora halophila and a related isolate have been shown to produce betaine from simple carbon sources (Severin et al., 1992; Galinski, 1993). There are few examples of other organisms which have been shown to be able to synthesize betaine from simple carbon sources. The data is usually based on metabolic studies using NMR. Roberts and co-workers (1992) have shown that some archaebacterial methanogens (Methanohalophilus) could synthesize betaine from glycine via a methylation reaction. However, no enzymes catalyzing the reactions have successfully been isolated from these organisms. A similar pathway has been suggested to exist in Ectothiorhodospira halochloris (Galinski and Trüper, 1994).

Attempts to characterize the glycine biosynthesis pathway in Ectothiorhodospira halochloris have been made (Tschichholtz-Mikus, 1994). According to the hypothesis proposed by this author, betaine would be synthesized from glycine by three methyltransferases, each specific for one methylation reaction. The isolation of the purified enzymes was, however, not successful and only one enzyme, specified as dimethyl glycine methyltransferase was partially purified. In addition, the methodology used to study reactions was rather simple and as demonstrated herein, the results obtained by the group differ from those obtained by the present inventors.

Betaine is used as feed additive in feed industry. Thus, transgenic plants producing high amounts of betaine in vivo would have better nutritional value. Feed crops (e.g. maize or soybean) producing sufficient amounts of betaine could therefore directly be used in feed without the need of betaine supplementation.

Although betaine is synthesized by many plants, there are several commercially important crops such as potato, rice, tomato and tobacco which do not accumulate betaine. For example, Bulow and co-workers (1995) were the first to demonstrate that expression of the E. coli choline dehydrogenase in tobacco improves the salt tolerance and freezing tolerance (Holmberg, 1996) of transgenic potato and tobacco due to endogenously synthesized betaine. The same phenomenon has been demonstrated also with Arabidopsis (Hayashi et al., 1997) and rice (Nakamura et at., 1997; Guo, 1997). Therefore, expression of the methyltransferases in plants can facilitate stress tolerance and improve the productivity of the plants when grown under conditions of water stress or freezing and cold temperatures.

Moreover, betaine has shown to induce pathogenesis-related protein expression in plants (Xin et al., 1996) as well as increasing the resistance of plants to attack by pathogenic

fungi or nematodes (Blunden et al., 1996; Wu et al. ,1997) and may decrease the incidence of nematode (e.g. Meloidogne javanica and M. incognita) attack in plants. Therefore, transgenic plants producing endogenous betaine, can be more resistant to fungal pathogens. Moreover, the betaine synthesis in the plants may be coupled to the systemic resistance genes which are induced when the plant is attacked by pathogens.

Endogenously synthesized betaine may also affect the viability of microbes and therefore it would improve their performance in various biotechnical processes. For instance, in high cell density fermentation or immobilized cell systems, the production microbes are subjected to considerable environmental stress. Betaine has successfully been used in fermentation media to increase the product yield in amino acid production. For instance, betaine has shown to relieve stress and improve yield of lysine producing Brevibacterium lactofermentum (Kawahara et al., 1990). Betaine is also commercially sold for the purpose (Nutristim®, Cultor Corp). Thus, endogenously synthesized betaine can improve productivity in biotechnical processes where the cells are subjected to abiotic stress.

Microbes also suffer from stress when subjected to high temperatures or when cells are freeze-dried or frozen. The viability of yeast or bacterial cells may be dramatically reduced in these processes used in e.g. frozen dough manufacturing or preservation of lactic acid bacterium starters. Therefore, it would be highly advantageous if one could improve the viability of microbes subjected to freeze-thaw or freeze-drying processes by accumulating betaine inside the cells. In addition, it has been shown that exogenously applied betaine improves the viability of microbes in extreme pH (Smirnova and Oktyabrsky, 1995; Chambers and Kunin, 1985).

Improved performance of beneficial, probiotic microbes organisms in animal digestive tract can be utilized in animal nutrition. Thus, introduction of a betaine synthesis pathway can improve the stress tolerance of "probiotic" lactic acid microbes which efficiently bind to gut epithelium in cells, providing a way to balance the microbial population in the GI tract and to improve pathogen resistance.

Betaine has been shown to stabilize proteins in the cells. For example, it has also been demonstrated that cytoplasmic accumulation of betaine will reduce the formation of inclusion bodies (Blackwell and Horgan, 1991; Bhandari and Gowrishankar, 1997) which is a problem often encountered when heterologous proteins are expressed in E. coli. Thus, co-expression of the genes of betaine biosynthesis with the protein of interest should result in better solubility of the heterologous protein and reduce the amount of inclusion bodies.

The use of the glycine methylation pathway may have a number of advantages over the oxidative synthesis from choline. Glycine is synthesized by practically all organisms and as an amino acid, this metabolite is present in high concentrations in the cells. In contrast, the availability of intracellular choline may limit betaine biosynthesis. In addition, the metabolism and formation of glycine in cells is known and the genes of this metabolic pathway have been cloned, thus allowing the engineering of the glycine pathway.

Based on the above, an object of the present invention is to provide proteins which are capable of acting in a biosynthetic pathway from glycine to betaine as well as methods for the purification and production of said proteins.

A further object of the present invention is to provide nucleic acid molecules which, when transformed into a host organism, encode proteins which are capable of acting in a biosynthetic pathway from glycine to betaine.

A further object is to provide recombinant microorganisms which are capable of expressing one or more proteins which are capable of acting in a biosynthetic pathway from glycine to betaine for the production of betaine and precursors thereof.

Furthermore, an object of the present invention is to provide recombinant plants which express one or more proteins which are capable of acting in a biosynthetic pathway from glycine to betaine for the production of betaine and precursors thereof.

Furthermore, an object of the present invention is to provide a method for the production of betaine and precursors thereof, for example sarcosine and dimethyl glycine, in recombinant organisms.

A further object is to provide recombinant organisms which have an increased concentration of intracellular betaine and are useful in the fields of recombinant heterologous protein production, agriculture, etc.

A further object of the present invention is to provide nucleic acid probes and method for identifying and cloning genes which encode proteins which are capable of participating in a biosynthetic pathway from glycine to betaine.

A further object of the present invention is to provide methods for improving the general growth and/or productivity of an organism including enhancing stress tolerance, for example, salt tolerance, freezing tolerance and cold tolerance, enhancing resistance to drought, water stress and attack by pathogens in organisms.

Furthermore, an object of the present invention is to provide recombinant microorganisms which have improved viability in

culture, enhanced pH tolerance in culture, result in decreased inclusion body formation when expressing a heterologous protein, result in increased solubility, stability and/or yield of a heterologous protein expressed in said organism.

Moreover, it is an object of the present invention to provide an animal feed and animal feed ingredient having enhanced nutritional value.

Other objects of the present invention will be apparent to the skilled person based on the information provided herein.

Summary of the Invention

The inventors have identified, isolated and purified proteins which are capable of carrying out at least one of the following reactions in a metabolic pathway from glycine to betaine: the conversion of glycine to sarcosine (N-methylglycine), sarcosine to N,N-dimethyl glycine and N,N-dimethyl glycine to betaine (N,N,N-trimethylglycine) in the presence of a methyl group donor. These proteins are designated herein as methyltransferases based on their ability to transfer a methyl group from a methyl group donor to a methyl group acceptor.

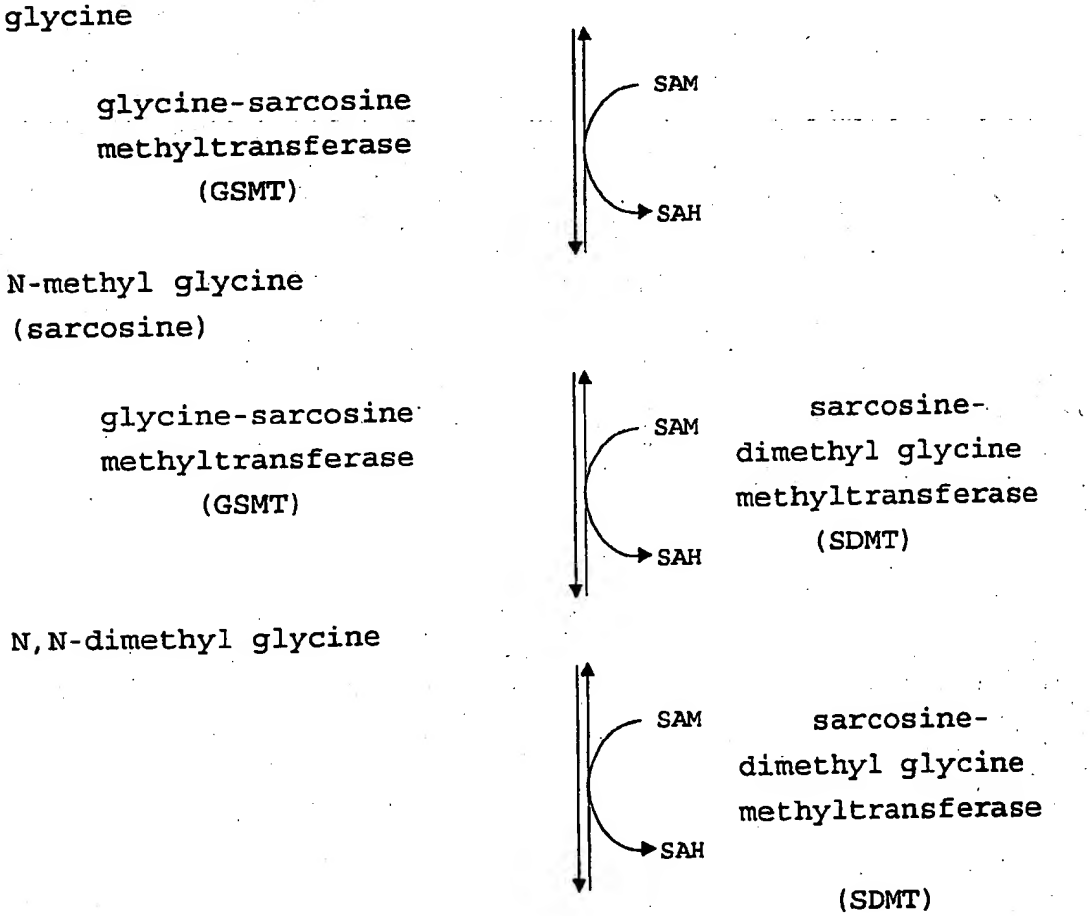
The above mentioned reactions are individual steps in a three-step methylation reaction pathway of glycine to betaine in certain microorganisms, for example, Ectothiorhodospira halochloris and Actinopolyspora halophila.

For example, a methyltransferase of the present invention (designated hereinafter as glycine-sarcosine methyltransferase or GSMT) has been isolated from Ectothiorhodospira halochloris and Actinopolyspora halophila which is capable of catalyzing methylation reactions that convert glycine to dimethyl glycine, i.e. convert glycine to

sarcosine (N-methyl glycine) and sarcosine to dimethyl glycine (N,N-dimethyl glycine).

As a further example, a methyltransferase according to the present invention (designated hereinafter as sarcosine-dimethylglycine methyltransferase or SDMT) has been isolated from Ectothiorhodospira halochloris and Actinopolyspora halophila which is capable of catalyzing methylation reactions that convert sarcosine to betaine, i.e. convert sarcosine (N-methyl glycine) to dimethyl glycine (N,N-dimethyl glycine) and dimethyl glycine to betaine.

The activities of GSMT and SDMT are described below.



The methyltransferases of the present invention are capable of utilizing S-adenosyl methionine (hereinafter also referred to as SAM) as a methyl group donor in the above reactions.

Brief Description of the Figures

Figure 1. Formation of methylation products from glycine, sarcosine and dimethyl glycine substrates using the *A. halophila* cell extract. The retention times of the standards are shown by arrows.

Figure 2. Analysis of purified methyl transferases on SDS-PAGE. A) *E. halochloris* GSMT; B) *A. halophila* SDMT. Lane 1, Purified protein sample; lane 2, molecular weight marker.

Figure 3. The determination of the isoelectric point of *A. halophila* SDMT by isoelectric focusing.

Figure 4. The pH-optimum of *A. halophila* SDMT. (●) Activity on sarcosine; (♠) Activity on dimethyl glycine

Figure 5. The temperature dependence of *A. halophila* SDMT activity. (●) Activity on sarcosine; (♠) Activity on dimethyl glycine.

Figure 6. *In vitro* synthesis of betaine by using the purified *E. halochloris* GSMT and *A. halophila* SDMT enzymes. The retention times of the standards are shown by arrows.

Figure 7. The schematic structure of the betaine operons of *A. halophila* and *E. halochloris*. GSMT; glycine sarcosine methyltransferase. SDMT; sarcosine dimethyl glycine methyltransferase. SAMS; S-adenosyl methionine synthase.

Figure 8. The nucleotide and amino acid sequence of the *E. halochloris* betaine operon. The arrows indicate the amino acids encoding GSMT, SDMT and SAMS. The underlined regions

indicate regions which are hybridized with the primers used to construct the expression vectors in heterologous organisms. The * indicates a stop codon.

Figure 9. The nucleotide and amino acid sequence of the *A. halophila* betaine operon. The underlined regions indicate regions which are hybridized with the primers used to construct the expression vectors in heterologous organisms. The * indicates a stop codon.

Figure 10. Schematic presentation of the expression plasmid used in expression of the methyl transferases. The insert was ligated to vector digested with *NcoI/BglII*.

Figure 11. The growth curves of *E. coli* transformants carrying the *E. halochloris* GSMT gene (EGSM). Transformant carrying only the cloning vector (PQE-60) was used as the control.

Figure 12. The growth curves of *E. coli* transformants carrying the *E. halochloris* GSMT and SDMT genes (EhFU). Transformant carrying only the cloning vector (PQE-60) was used as the control.

Detailed Description of the Invention

One embodiment of the present invention provides a methyltransferase, for example GSMT, capable of catalyzing the conversion of glycine to sarcosine (N-methyl glycine) and/or the conversion of sarcosine to dimethyl glycine (N,N-dimethyl glycine). Preferably, said methyltransferase comprises an amino acid sequence selected from the group consisting of:

(a) an amino acid sequence as depicted in SEQ ID NO:2 and an amino acid sequence as depicted in SEQ ID NO:6,

(b) a fragment of an amino acid sequence as defined in (a) and

(c) a derivative of an amino acid sequence as defined in (a) and (b).

A fragment of an amino acid sequence as depicted in SEQ ID NO:2 or SEQ ID NO:6 is designated as any fragment of an amino acid sequence as depicted in SEQ ID NO:2 or SEQ ID NO:6 which is capable of catalyzing the conversion of glycine to sarcosine (N-methyl glycine) and/or the conversion of sarcosine to dimethyl glycine (N,N-dimethyl glycine).

A derivative of an amino acid sequence as depicted in SEQ ID NO:2 or SEQ ID NO:6 is designated as any mutation, deletion, addition, substitution, insertion or inversion of one or more amino acids in the of amino acid sequence as depicted in SEQ ID NO:2 or SEQ ID NO:6 which is capable of catalyzing the conversion of glycine to sarcosine (N-methyl glycine) and/or the conversion of sarcosine to dimethyl glycine (N,N-dimethyl glycine). Preferably, a derivative of a methyltransferase which is capable of catalyzing the conversion of glycine to sarcosine (N-methyl glycine) and/or the conversion of sarcosine to dimethyl glycine (N,N-dimethyl glycine) has about 60 % homology, preferably about 70 % homology, more preferably about 80 % homology, and most preferably about 90 % homology to the corresponding amino acid sequence depicted in SEQ ID NO:2 or SEQ ID NO:6. Preferably, when the above amino acid sequences has one or more mutations, substitutions, additions and/or insertions, the amino acids constituting these changes are selected from the 20 standard naturally-occurring amino acids found in proteins, i.e. Ala, Val, Leu, Ile, Pro, Phe, Trp, Met, Gly, Ser, Thr, Cys, Tyr, Asn, Gln, Asp, Glu, Lys, Arg and His. Preferably, the mutation(s) and/or substitution(s) are conservative, for example, Ala, Val, Leu, Ile, Pro, Phe, Trp, or Met residue(s) are replaced with one of these amino acids, Gly, Ser, Thr, Cys, Tyr, Asn, Gln, residue(s) are replaced with one of these amino acids, Asp or Glu are replaced with one of these amino acids and Lys, Arg or His are replaced with one of these amino acids.

In a preferred embodiment, a methyltransferase according to the invention has the amino acid sequence depicted in SEQ ID NO:2.

In another preferred embodiment, a methyltransferase according to the invention has the amino acid sequence depicted in SEQ ID NO:6.

The methyltransferase of the present invention capable of catalyzing the conversion of glycine to sarcosine (N-methyl glycine) and/or the conversion of sarcosine to dimethyl glycine (N,N-dimethyl glycine) can exist in the form of an active enzyme or as a zymogen. The term 'zymogen' designates a protein molecule or fragment or derivative thereof (as defined above) which is synthesized in an inactive form and is capable of being activated in vitro or in vivo by the chemical or enzymatic cleavage of one or more peptide bonds. A preferred zymogen of the methyltransferase of the present invention capable of catalyzing the conversion of glycine to sarcosine (N-methyl glycine) and/or the conversion of sarcosine to dimethyl glycine (N,N-dimethyl glycine) comprises the amino acid sequence as depicted in SEQ ID NO:2 and SEQ ID NO: 3, wherein the N-terminus of SEQ ID NO: 3 is joined to the C-terminus of SEQ ID NO:2.

A further embodiment of the present invention provides a methyltransferase, for example SDMT, capable of catalyzing the conversion of sarcosine (N-methyl glycine) to dimethyl glycine (N,N-dimethyl glycine) and/or dimethyl glycine to betaine. Preferably, said methyltransferase comprises an amino acid sequence selected from the group consisting of:

(a) an amino acid sequence as depicted in SEQ ID NO:3 and an amino acid sequence as depicted in SEQ ID NO:7,

(b) a fragment of an amino acid sequence as defined in (a) and

(c) a derivative of an amino acid sequence as defined in (a) and (b).

A fragment of an amino acid sequence as depicted in SEQ ID NO:3 or SEQ ID NO:7 is designated as any fragment of an amino acid sequence as depicted in SEQ ID NO:3 or SEQ ID NO:7 which is capable of catalyzing the conversion of sarcosine (N-methyl glycine) to dimethyl glycine and/or the conversion of dimethyl glycine to betaine.

A derivative of an amino acid sequence as depicted in SEQ ID NO:3 or SEQ ID NO:7 is designated as any mutation, deletion, addition, substitution, insertion or inversion of one or more amino acids, or combination thereof, in the of amino acid sequence as depicted in SEQ ID NO:3 or SEQ ID NO:7 which is capable of catalyzing the conversion of sarcosine (N-methyl glycine) to dimethyl glycine and/or the conversion of dimethyl glycine (N,N-dimethyl glycine) to betaine. Preferably, a derivative of a methyltransferase which is capable of catalyzing the conversion of sarcosine (N-methyl glycine) to dimethyl glycine and/or the conversion of dimethyl glycine (N,N-dimethyl glycine) to betaine has about 60 % homology, preferably about 70 % homology, more preferably about 80 % homology, and most preferably about 90 % homology to the corresponding amino acid sequence depicted in SEQ ID NO:3 or SEQ ID NO:7. Preferably, when the above amino acid sequences has one or more mutations, substitutions, additions and/or insertions, the amino acids constituting these changes are selected from the 20 standard naturally-occurring amino acids found in proteins, i.e. Ala, Val, Leu, Ile, Pro, Phe, Trp, Met, Gly, Ser, Thr, Cys, Tyr, Asn, Gln, Asp, Glu, Lys, Arg and His. Preferably, the mutation(s) and/or substitution(s) are conservative, for example, Ala, Val, Leu, Ile, Pro, Phe, Trp, or Met residue(s) are replaced with one of these amino acids, Gly, Ser, Thr, Cys, Tyr, Asn, Gln, residue(s) are replaced with one of these amino acids, Asp or Glu are replaced with one of these amino acids and Lys, Arg or His are replaced with one of these amino acids.

In a preferred embodiment, a methyltransferase according to the invention has the amino acid sequence depicted in SEQ ID NO:3.

In another preferred embodiment, a methyltransferase according to the invention has the amino acid sequence depicted in SEQ ID NO:7.

The methyltransferase of the present invention capable of catalyzing the conversion of sarcosine (N-methyl glycine) to dimethyl glycine and/or the conversion of dimethyl glycine (N,N-dimethyl glycine) to betaine can exist in the form of an active enzyme or as a zymogen. The term 'zymogen' designates a protein molecule, fragment or derivative thereof (as defined above) which is synthesized in an inactive form and is capable of being activated in vitro or in vivo by the chemical or enzymatic cleavage of one or more peptide bonds. A preferred zymogen of the methyltransferase of the present invention capable of catalyzing the conversion of sarcosine (N-methyl glycine) to dimethyl glycine and/or the conversion of dimethyl glycine (N,N-dimethyl glycine) to betaine comprises the amino acid sequence as depicted in SEQ ID NO:2 and SEQ ID NO: 3, wherein the N-terminus of SEQ ID NO: 3 is joined to the C-terminus of SEQ ID NO:2.

The methyltransferases of the present invention, for example naturally occurring GSMT and SDMT or recombinantly produced GSMT and SDMT and fragments or derivatives thereof as well as zymogens of these naturally occurring or recombinant proteins, are preferably isolated to a state free from other proteins originating from the organisms from which they are isolated; more preferably, to a pure state, most preferably to a homogeneous state.

Another aspect of the present invention provides an enzyme capable of catalyzing the synthesis of S-adenosyl methionine (SAM), i.e. S-adenosyl methionine synthase (hereinafter referred to as SAMS) which converts methionine to S-adenosyl

methionine in the presence of ATP. Preferably, said SAMS comprises an amino acid sequence selected from the group consisting of:

(a) an amino acid sequence as depicted in SEQ ID NO:4 and an amino acid sequence as depicted in SEQ ID NO:8,

(b) a fragment of an amino acid sequence as defined in (a) and

(c) a derivative of an amino acid sequence as defined in (a) and (b).

A fragment of an amino acid sequence as depicted in SEQ ID NO:4 or SEQ ID NO:8 is designated as any fragment of an amino acid sequence as depicted in SEQ ID NO:4 or SEQ ID NO:8 which is capable of catalyzing the conversion of S-adenosyl methionine from methionine and ATP (adenosine triphosphate).

A derivative of an amino acid sequence as depicted in SEQ ID NO:4 or SEQ ID NO:8 is designated as any mutation, deletion, addition, substitution, insertion or inversion of one or more amino acids, or combination thereof, in the of amino acid sequence as depicted in SEQ ID NO:4 or SEQ ID NO:8 which is capable of catalyzing the conversion of S-adenosyl methionine from methionine and ATP. Preferably, a derivative which is capable of catalyzing the conversion of methionine to S-adenosyl methionine has about 60 % homology, preferably about 70 % homology, more preferably about 80 % homology, and most preferably about 90 % homology to the corresponding amino acid sequence depicted in SEQ ID NO:4 or SEQ ID NO:8. Preferably, when the above amino acid sequences has one or more mutations, substitutions, additions and/or insertions, the amino acids constituting these changes are selected from the 20 standard naturally-occurring amino acids found in proteins, i.e. Ala, Val, Leu, Ile, Pro, Phe, Trp, Met, Gly, Ser, Thr, Cys, Tyr, Asn, Gln, Asp, Glu, Lys, Arg and His. Preferably, the mutation(s) and/or substitution(s) are conservative, for example, Ala, Val, Leu, Ile, Pro, Phe, Trp, or Met residue(s) are replaced with one of these amino acids, Gly, Ser, Thr, Cys, Tyr, Asn, Gln, residue(s) are replac d

with one of these amino acids, Asp or Glu are replaced with one of these amino acids and Lys, Arg or His are replaced with one of these amino acids.

In a preferred embodiment, the SAMS of the invention has the amino acid sequence depicted in SEQ ID NO:4.

In another preferred embodiment, the SAMS of the invention has the amino acid sequence depicted in SEQ ID NO:8.

The SAMS of the present invention, for example naturally occurring SAMS or recombinantly produced SAMS and fragments or derivatives thereof, are preferably isolated to a state free from other proteins originating from the organisms from which they are isolated; more preferably, to a pure state, most preferably to a homogeneous state.

The present invention also relates to a nucleic acid molecule which is capable of encoding a methyltransferase capable of converting glycine to sarcosine (N-methyl glycine) and/or sarcosine to dimethyl glycine (N,N-dimethyl glycine).

In a preferred embodiment of the present invention, a nucleic acid molecule which encodes a methyltransferase capable of converting glycine to sarcosine (N-methyl glycine) and/or sarcosine to dimethyl glycine (N,N-dimethyl glycine) comprises a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1 and

a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

A nucleotide sequence which hybridizes to a nucleic acid molecule comprising a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1 or a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5 is designated as any nucleotide sequence, for example DNA or RNA, preferably DNA, which hybridizes under standard conditions to a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1 or a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5 and encodes a methyltransferase capable of converting glycine to sarcosine (N-methyl glycine) and/or sarcosine to dimethyl glycine (N,N-dimethyl glycine). The term 'standard conditions' designates a standard procedure used in heterologous hybridization to screen for genes with enough sequence homology (Maniatis, 1989). Basically, the hybridization filters are washed first in low stringency conditions and the stringency is gradually increased (usually by increasing the washing temperature) in order to select the positive signals from the background. For example, plaques or colonies of a gene bank to be screened can be transferred onto nitrocellulose membranes and hybridized with a PCR fragment, oligonucleotide or any other cloned DNA containing a fragment of the above gene of interest. The probe can be prepared for example by PCR as described in Example 6. Hybridization can be carried out at 42°C in 50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 1 M NaCl, 0.5% SDS, 0.1% sodium pyrophosphate, 10x Denhardt's solution (Maniatis, 1989), 100 µg herring sperm DNA and 125 µg/ml polyA. The filters can be first washed with low stringency conditions, for example, at 37°C using 3x SSC, 0.5% SDS and 10% sodium pyrophosphate (Maniatis, 1989). The filters can then be exposed to x-ray film to monitor the number of positive clones. The washing temperature will be raised in 2-5°C intervals up to a temperature at which the background becomes invisible and only a few positive plaques are obtained. The positive plaques or colonies can then be

purified and the DNA can be isolated for Southern blot hybridization to check the size of the cloned insert for example. The cloned DNA obtained can be sequenced. On the basis of the sequence homology, it can be concluded that the DNA contains the gene of interest.

Preferably, a nucleotide sequence which hybridizes to a nucleic acid molecule comprising a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1 or a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5 has about 60 %, preferably 70 %, more preferably 80 % and especially 90 % homology to a nucleotide sequence corresponding to a DNA sequence comprising nucleotide 208 to 1047 of SEQ ID NO:1 or a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5.

A fragment of a nucleotide sequence from nucleotide 208 to 1047 of SEQ ID NO:1 or a nucleotide sequence from nucleotide 221 to 1024 of SEQ ID NO:5 is designated as any nucleic acid fragment, for example DNA or RNA, preferably DNA, which encodes a methyltransferase capable of converting glycine to sarcosine (N-methyl glycine) and/or sarcosine to dimethyl glycine (N,N-dimethyl glycine).

A preferred nucleic acid molecule according to the invention comprises the DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1.

A further preferred nucleic acid molecule according to the invention comprises the DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5

Other preferred nucleic acid molecules according to the invention comprise the DNA sequence from nucleotide 221 to 1867 of SEQ ID NO:5 or the DNA sequence from nucleotide 208 to 1902 of SEQ ID NO:1, the DNA sequence from nucleotide 221 to 1867 of SEQ ID NO:5 being more preferred.

The present invention also relates to a nucleic acid molecule which encodes a methyltransferase capable of converting sarcosine (N-methyl glycine) to dimethyl glycine and/or dimethyl glycine (N,N-dimethyl glycine) to betaine.

In a preferred embodiment of the present invention, a nucleic acid molecule which encodes a methyltransferase capable of converting sarcosine (N-methyl glycine) to dimethyl glycine and/or dimethyl glycine (N,N-dimethyl glycine) to betaine comprises a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1 and

a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

A nucleotide sequence which hybridizes to a nucleic acid molecule comprising a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1 or a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5 is designated as any nucleotide sequence, for example DNA or RNA, preferably DNA, which hybridizes under standard conditions to a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1 or a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5 and encodes a methyltransferase capable of converting sarcosine (N-methyl glycine) to dimethyl glycine and/or dimethyl glycine (N,N-dimethyl glycine) to betaine. The term 'standard conditions' designates a standard procedure used in heterologous hybridization to screen for genes with enough sequence homology (Maniatis, 1989). Basically, the hybridization filters are washed first in low stringency conditions and the

stringency is gradually increased (usually by increasing the washing temperature) in order to select the positive signals from the background. For example, plaques or colonies of a gene bank to be screened can be transferred onto nitrocellulose membranes and hybridized with a PCR fragment, oligonucleotide or any other cloned DNA containing a fragment of the above gene of interest. The probe can be prepared for example by PCR as described in Example 6. Hybridization can be carried out at 42°C in 50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 1 M NaCl, 0.5% SDS, 0.1% sodium pyrophosphate, 10x Denhardt's solution (Maniatis, 1989), 100 µg herring sperm DNA and 125 µg/ml polyA. The filters can be first washed with low stringency conditions, for example, at 37°C using 3x SSC, 0.5% SDS and 10% sodium pyrophosphate (Maniatis, 1989). The filters can then be exposed to x-ray film to monitor the number of positive clones. The washing temperature will be raised in 2-5°C intervals up to a temperature at which the background becomes invisible and only a few positive plaques are obtained. The positive plaques or colonies can then be purified and the DNA can be isolated for Southern blot hybridization to check the size of the cloned insert for example. The cloned DNA obtained can be sequenced. On the basis of the sequence homology, it can be concluded that the DNA contains the gene of interest.

Preferably, a nucleotide sequence which hybridizes to a nucleic acid molecule comprising a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1 or a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5 has about 60 %, preferably 70 %, more preferably 80 % and especially 90 % homology to a nucleotide sequence corresponding to a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1 or a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5.

A fragment of a nucleotide sequence from nucleotide 1048 to 1902 of SEQ ID NO:1 or a nucleotide sequence from nucleotide 1031 to 1867 of SEQ ID NO:5 is designated as any nucleic acid fragment, for example DNA or RNA, preferably DNA, which

encodes a methyltransferase capable of converting sarcosine (N-methyl glycine) to dimethyl glycine and/or dimethyl glycine (N,N-dimethyl glycine to glycine betaine.

A preferred nucleic acid molecule according to the invention comprises the DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1.

A further preferred nucleic acid molecule according to the invention comprises the DNA sequence from 1031 to 1867 of SEQ ID NO:5.

Other preferred nucleic acid molecules according to the invention comprise the DNA sequence from nucleotide 221 to 1867 of SEQ ID NO:5 or the DNA sequence from nucleotide 208 to 1902 of SEQ ID NO:1, the DNA sequence from nucleotide 221 to 1867 of SEQ ID NO:5 being more preferred.

A further aspect of the present invention provides a nucleic acid molecule which encodes an enzyme capable of converting S-adenosyl methionine from methionine and ATP.

In a preferred embodiment of the nuclic acid molecule provided by the present invention, a nucleic acid molecule which encodes an enzyme capable of converting S-adenosyl methionine from methionine and ATP (SAMS) comprises a nucleotide sequence selected from the group consisting of:
(a) a DNA sequence from nucleotide 2027 to 2722 of SEQ ID NO:1 and
a DNA sequence from nucleotide 2006 to 3004 of SEQ ID NO:5,
(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),
(c) a fragment of a nucleotide sequence as defined in (a) and (b), and
(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c).

A nucleotide sequence which hybridizes to a nucleic acid molecule comprising a DNA sequence from nucleotide 2027 to 2722 of SEQ ID NO:1 or a DNA sequence from nucleotide 2006 to 3004 of SEQ ID NO:5 is designated as any nucleotide sequence, for example DNA or RNA, preferably DNA, which hybridizes under standard conditions to a DNA sequence from nucleotide 2027 to 2722 of SEQ ID NO:1 or a DNA sequence from nucleotide 2006 to 3004 of SEQ ID NO:5 and encodes enzyme capable of converting S-adenosyl methionine from methionine and ATP. The term 'standard conditions' designates a standard procedure used in heterologous hybridization to screen for genes with enough sequence homology (Maniatis, 1989). Basically, the hybridization filters are washed first in low stringency conditions and the stringency is gradually increased (usually by increasing the washing temperature) in order to select the positive signals from the background. For example, plaques or colonies of a gene bank to be screened can be transferred onto nitrocellulose membranes and hybridized with a PCR fragment, oligonucleotide or any other cloned DNA containing a fragment of the above gene of interest. The probe can be prepared for example by PCR as described in Example 6. Hybridization can be carried out at 42°C in 50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 1 M NaCl, 0.5% SDS, 0.1% sodium pyrophosphate, 10x Denhardt's solution (Maniatis, 1989), 100 µg herring sperm DNA and 125 µg/ml polyA. The filters can be first washed with low stringency conditions, for example, at 37°C using 3x SSC, 0.5% SDS and 10% sodium pyrophosphate (Maniatis, 1989). The filters can then be exposed to x-ray film to monitor the number of positive clones. The washing temperature will be raised in 2-5°C intervals up to a temperature at which the background becomes invisible and only a few positive plaques are obtained. The positive plaques or colonies can then be purified and the DNA can be isolated for Southern blot hybridization to check the size of the cloned insert for example. The cloned DNA obtained can be sequenced. On the basis of the sequence homology, it can be concluded that the DNA contains the gene of interest.

Preferably, a nucleotide sequence which hybridizes to a nucleic acid molecule comprising a DNA sequence from nucleotide 2027 to 2722 of SEQ ID NO:1 or a DNA sequence from nucleotide 2006 to 3004 of SEQ ID NO:5 has about 60 %, preferably 70 %, more preferably 80 % and especially 90 % homology to a nucleotide sequence corresponding to a DNA sequence from nucleotide 2027 to 2722 of SEQ ID NO:1 or a DNA sequence from nucleotide 2006 to 3004 of SEQ ID NO:5.

A fragment of a nucleotide sequence from nucleotide 2027 to 2722 of SEQ ID NO:1 or a DNA sequence from nucleotide 2006 to 3004 of SEQ ID NO:5 is designated as any nucleic acid fragment, for example DNA or RNA, preferably DNA, which encodes enzyme capable of converting S-adenosyl methionine from methionine and ATP.

A preferred nucleic acid molecule according to the invention comprises the DNA sequence from nucleotide 2027 to 2722 of SEQ ID NO:1.

A further preferred nucleic acid molecule according to the invention comprises the DNA sequence from nucleotide 2006 to 3004 of SEQ ID NO:5.

Further subject matter of the invention is a DNA probe for use in identifying and cloning a nucleic acid molecule encoding a methyltransferase comprising at least 15 nucleotide bases, preferably 20 or more nucleotide bases, of a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,

a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,

a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,

a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5.

Said DNA probes can be utilized in a method according to the invention for identifying and cloning a nucleic acid molecule encoding a methyltransferase comprising the steps of hybridizing said probe with a sample containing nucleic acid of an organism, detecting a nucleic acid molecule in said sample which hybridizes to said probe and isolating said detected nucleic acid molecule. Preferred methods include the use of the polymerase chain reaction (PCR) and Southern blotting techniques which are described herein and are familiar to the skilled person in the art.

Further subject matter of the invention are vectors for expression of the proteins according to the invention in prokaryotic and eukaryotic hosts.

In this connection, expression vectors, for example phages, plasmids and DNA or RNA viruses, are capable of transforming and/or replicating and expressing the proteins of the present invention in prokaryotes and/or eukaryotes, for example bacteria, yeast, fungi and/or plants. Such expression vectors and methods for their construction are known to the skilled person and can be provided with nucleic acid elements for transcription, for example start codons, 'TATA' boxes, promoters, enhancers, stop codons, etc., and nucleic acid elements important for translation and processing of the nucleic acids transcribed from said vectors in a given host, for example ribosome binding sites, leader sequences for secretion of the proteins of the present invention, etc.

One embodiment of the invention is an expression vector comprising a nucleic acid sequence which encodes a methyltransferase capable of converting glycine to sarcosine (N-methyl glycine) and/or sarcosine to dimethyl glycine (N,N-dimethyl glycine) and/or a nucleic acid sequence which encodes a methyltransferase capable of converting sarcosine (N-methyl glycine) to dimethyl glycine and/or dimethyl glycine (N,N-dimethyl glycine) to betaine.

In a preferred embodiment, said expression vector comprises at least one nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,

a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,

a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,

a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c).

In a further preferred embodiment, an expression vector is provided comprising a nucleotide sequence coding for an enzyme capable of catalyzing the synthesis of S-adenosyl methionine and at least one nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,

a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,

a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,

a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

When an expression vector of the present invention contains one of the above mentioned DNA sequences, preferred expression vectors comprise a nucleotide sequence selected from the group consisting of:

(a) the DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,

the DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,

the DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1, or

the DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5.

When an expression vector of the present invention contains two of the above mentioned DNA sequences, preferred expression vectors comprise the DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1 and the DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1 or the DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5 and the DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5, an expression vector comprising the DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5 and the DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5 being more preferred.

However, expression vectors comprising fragments and/or derivatives of the above mentioned sequences as well as other combinations of the above mentioned DNA sequences, for example a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1 and a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5 or a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5 and a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1 or fragments and/or derivatives thereof are also subject matter of the present invention.

As provided for above, expression vectors of the present invention can additionally comprise a nucleic acid molecule coding for an enzyme capable of directly or indirectly increasing the intracellular amount of S-adenosyl methionine.

The genes encoding E. halochloris and A. halophila GSMT and SDMT enzymes are located in a "betaine operon". In E. halochloris the enzymes are encoded by two separate genes, whereas in A. halophila the two enzymes are coded by a single gene. In addition, the "betaine operon" contains a S-adenosyl methionine synthase (SAMS) gene. The SAMS enzyme catalyzes the synthesis of S-adenosyl methionine (SAM) from methionine and ATP, and thus, it is useful in the methylation reactions of the methyltransferases of the invention because it increases the concentration of the enzyme substrate SAM. Therefore co-expression of the SAMS gene with one or more of the methyltransferase genes of the invention can be used to increase betaine synthesis in these organisms.

Hence, in a preferred embodiment, the above mentioned expression vectors can additionally comprise a DNA sequence from nucleotide 2027 to 2722 of SEQ ID NO:1 or a DNA sequence from nucleotide 2006 to 3004 of SEQ ID NO:5 or fragments and/or derivatives thereof when the organism to be transformed is E halochloris or A. halophila.

Preferred expression vectors of this type, which also encode the methyltransferases of the present invention, comprise a DNA sequence from nucleotide 208 to 2722 of SEQ ID NO:1 or a DNA sequence from nucleotide 221 to 3004 of SEQ ID NO:5, the latter being more preferred.

Alternatively, said nucleic acid molecule coding for an enzyme capable of directly or indirectly increasing the intracellular amount of S-adenosyl methionine originates or is derived from the organism which is to be transformed with said expression vector. For example, if the organism to be transformed with a nucleic acid, for example an expression

vector, according to the invention is E. coli, then it is possible for example to incorporate the gene coding for S-adenosyl methionine synthase from E. coli (see Markham et al., 1984). In a similar manner, if the organism to be transformed with a nucleic acid, for example an expression vector, according to the invention is Bacillus, then it is possible for example to incorporate the gene coding for S-adenosyl methionine synthase from Bacillus subtilis (see Yocum et al., 1996, and genebank accession number AF008220). Other nucleotide sequences which can be used for this purpose are the SAH hydrolase form Mesembryanthemum crystallinum (genebank accession number U79766; Arabidopsis thaliana, accession number AF059581; S. pombe, accession number AL022072).

In another aspect of the present invention, the expression vectors can additionally comprise a nucleic acid molecule coding for an enzyme capable of increasing the intracellular amount of intracellular glycine. Preferably, said nucleic acid molecule coding for an enzyme capable of directly or indirectly increasing the intracellular amount of glycine originates or is derived from the organism which is to be transformed with said expression vector.

For example, the expression vector can include a nucleotide sequence encoding the enzyme phosphoglycerate dehydrogenase (Bacillus subtilis, accession number L47648; S. pombe, accession number AL022243; Arabidopsis thaliana, accession number AB010407), phosphoserine aminotransferase (E. coli, accession number AE000193, U00096; Bacillus subtilis, accession number Z99109, AL009126; S. pombe, accession number Z69944; Arabidopsis thaliana, accession number AL031135), phosphoserine phosphatase (E. coli, accession number AE000509; S. cerevisiae, accession number U36473; S. pombe, accession number D89261) and serine hydroxymethyl transferase (Bacillus thearothermophilius, E02190; Candida albicans, accession number AF009966; Zea mays, accession number W49449).

Further subject matter of the present invention is a recombinant prokaryotic or eukaryotic organism, for example, bacteria, yeast, fungus or plant, transformed with at least one nucleic acid molecule of the invention as defined above, for example, an expression vector according to the invention as defined above.

When a recombinant organism according to the invention is a bacterium, said bacterium is preferably selected from the group consisting of E. coli, Bacillus, Corynebacteria, Pseudomonas and lactic acid bacteria and Streptomyces.

Numerous genetic tools for expressing genes in lactic acid bacteria have been developed in the past few years. The field has extensively been reviewed by Kuipers et al. (1997). This opens up many possibilities to develop an inducible expression system for the GSMT or SDMT genes in lactic acid bacteria. In addition, the methodology to express heterologous genes in *Bacillus* is well established and there are number of functional expression systems facilitating the overexpression of the GSMT and SDMT in *Bacillus*. (Sarvas, M. (1994)).

When the recombinant organism according to the invention is a yeast, said yeast is preferably selected from the group consisting of Saccharomyces, Zygosaccharomyces, Pichia, Kluyveromyces, Candida and Hansenula.

When the recombinant organism according to the invention is a fungus, said fungus is preferably selected from the group consisting of Aspergillus, Trichoderma and Penicillium.

When the recombinant organism according to the invention is a plant including but not limited to cereals, legumes, oilseeds, vegetables, fruits, ornamentals and perennials, said plant is preferably selected from the group consisting of lettuces, Capsicums, grasses, clovers, alfalfa, beans,

sweet potatoes, cassava, yams, taro, groundnut, brassica, sugar beet, grapes, potato, tomato, rice, tobacco, rapeseed, maize, sorghum, cotton, soybean, barley, wheat, rye, canola, sunflower, linseed, pea, cucumber, carrot, ornamentals, perennial trees including citrus pear and almond and fruits including strawberry.

In this connection, the term 'plant' according to the invention is understood to include individual cells of a plant, plant seeds and callus material.

Further subject matter of the present invention is a method for the production of a recombinant organism according to the invention comprising the steps of transforming a host prokaryotic or eukaryotic organism, preferably a bacteria, yeast or fungus, with at least one nucleic acid molecule of the invention as defined above, for example an expression vector according to the invention as defined above. When the host organism is to be transformed with a nucleic acid molecule coding for an enzyme capable of increasing the intracellular amount of S-adenosyl methionine and/or glycine which originates from the host organism, this nucleic acid, for example an expression vector, can be transformed as a separate molecule or can be cloned into a expression vector according to the invention. Likewise, when the host organism is to be transformed with two different nucleic acid molecules each encoding a different methyltransferase of the invention, then transformation can be performed using these two nucleic acid molecules, e.g. expression vectors.

In addition, the present invention relates to a methyltransferase obtainable by culturing wild-type Ectothiorhodospira or Actinopolyspora or a recombinant prokaryotic or eukaryotic organism according to the invention and isolating said methyltransferase from the organism and/or the medium used to culture or process said organism as well as a method for the production of said methyltransferase comprising the above mentioned steps.

In this connection, a method for the purification of a methyltransferase capable of catalyzing the conversion of glycine to dimethyl glycine comprising the steps of subjecting a sample comprising the methyltransferase to a matrix containing adenosine, binding said methyltransferase to said matrix and eluting said methyltransferase from said matrix is also subject matter of the present invention. In addition, the above purification step can be combined with other methods of protein purification including ammonium sulfate precipitation, size exclusion chromatography, cation or anion exchange chromatography, hydrophobic interaction chromatography, etc.

By expressing the genes encoding the GSMT and SDMT enzymes, it is possible to impart the capability of de novo synthesis of betaine to different organisms. Suitable host organisms are practically all bacteria which can be transformed with foreign DNA (for instance E. coli, Bacillus, Corynebacteria, Pseudomonas, lactic acid bacteria and Streptomyces) yeast (for instance Saccharomyces, Zygosaccharomyces, Pichia, Kluyveromyces, Candida and Hansenula) fungi (for instance Trichoderma, Aspergillus, Penicillium) or plants.

Therefore, subject matter of the present invention is a method for the production of betaine comprising the steps of culturing a recombinant organism according to the invention and isolating betaine from the organism and/or the medium used to culture or process said organism.

Further subject matter of the present invention is a method for the production of sarcosine and/or dimethyl glycine comprising the steps of culturing an organism transformed with a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,

a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c),

and isolating sarcosine and/or dimethyl glycine from said organism or the medium used to culture or process said organism.

Further subject matter of the invention is a method for increasing the intracellular concentration of sarcosine, dimethyl glycine and/or betaine in an organism, enhancing the general productivity of an organism, enhancing the salt tolerance of an organism, enhancing the freezing or cold tolerance of an organism, and/or enhancing the resistance of an organism to drought and/or low water stress comprising the steps of transforming an organism with at least one nucleic acid molecule of the invention as defined above, for example an expression vector according to the invention as defined above. Preferably, said organism is a bacteria, yeast, fungus or plant as recited above.

Further subject matter of the invention is a method for inducing pathogenesis-related proteins in a plant, increasing the resistance of a plant to attack by pathogens and/or increasing the nutritional value of a plant comprising the steps of transforming a plant with at least one nucleic acid molecule of the invention as defined above, for example an expression vector according to the invention as defined above. Preferably, said plant is a plant as recited above.

Pathogens include but are not limited to Fusarium sp. which cases root, shoot and leaf diseases in several plant types, Rhizoctonia sp. and Pythium sp. which cause soil borne

diseases in crops, Erysiphe sp. which cause mildew in several species, Phytophthora infestans which causes late blight in potato and tomato, Alternaria solani which causes early blight in potato, fungal diseases of soya caused by Cephalosporium sp., Diaporthe sp., Cerospora sp. Septoria sp. and Peronospora sp., nematodes for example Meloidogne javanica and M. incognita and insects.

Moreover, subject matter of the invention is a method for enhancing the pH tolerance and/or viability of a cultured microorganism comprising the steps of transforming a microorganism with at least one nucleic acid molecule of the invention as defined above, for example an expression vector according to the invention as defined above. Preferably, said microorganism is a bacteria, yeast or fungus as recited above.

The microorganisms of the invention can also be used as hosts in the field of recombinant DNA technology for the expression of a heterologous protein of interest. Therefore, subject matter of the invention is a method for decreasing inclusion body formation, increasing the stability of a heterologous protein and/or increasing the production of a heterologous protein expressed in a microorganism comprising the steps of transforming a microorganism with at least one nucleic acid molecule of the invention as defined above, for example an expression vector according to the invention as defined above, and transforming a microorganism with a nucleic acid molecule capable of expressing said heterologous protein. Said microorganism can be transformed with the nucleic acid molecule, for example an expression vector, according to the invention before, during or after the microorganism is transformed with a nucleic acid molecule capable of expressing said heterologous protein.

Additional subject matter of the invention is an animal feed or animal feed ingredient comprising a recombinant organism according to the invention.

The present invention is more closely illustrated by means of the following examples without limiting the invention to the examples.

Examples

Example 1. Demonstration of the methyltransferase pathway in Actinopolyspora halophila and Ectothiorhodospira halochloris

Preparation of the cell extracts

The growth medium of Actinopolyspora halophila ATCC 27976 used in all cultivations was the "complex medium" described by Sehgal and Gibbons (1960). Inoculum was grown at 37°C in a shake flask with agitation at 180 rpm until the late exponential growth phase. Then, 8 l of the above medium with 10 g/l glucose was inoculated with 800 ml culture. The pH in the fermentor (Biostat M (Braun) laboratory fermentor) was maintained at pH 6.5-7.5 with 0.5 M H_2SO_4 and 1 M NaOH. Agitation and aeration rates were 400 rpm and 10 l/min., respectively. The cultivation temperature was 37°C. Cells were grown to late exponential phase and harvested by centrifugation at 15,000 g for 15 min. Cells were stored at -75°C. Before disruption, the cells were thawed and suspended in Buffer I (22 % (w/v) sucrose, 27 mM EDTA, 1 mM phenylmethylsulfonyl fluoride (PMSF), 50 mM Tris-HCl, pH 7.5) in a ratio of 1.5 ml buffer I:1 g cells (wet weight).

The growth medium used in the cultivation of Ectothiorhodospira halochloris ATCC 35916 is described by Tschichholz and Trüper (1990). Cultivation was carried out anaerobically at 42°C in 1 l glass bottles with continuous stirring with magnetic stirrer. The cells were illuminated during growth (5,000 - 10,000 lux). 100 ml of pre-inoculum was inoculated into 1 l medium. Cells were grown until late exponential phase and harvested by centrifugation at 28,000 g for 20 min. After centrifugation, the cells were suspended in

Buffer II (560 mM Tris-HCl, pH 7.5, 4 mM 2-mercaptoethanol, 50 µM $MgCl_2$, 160 µM EDTA) and disrupted with 1 mM PMSF and 1 mM dithiothreitol (DTT). Buffer II was added in a ratio of 1.5 ml buffer II: 1 g cells (wet weight).

The cells were disrupted with a MSE Soniprep 150 sonicator. The suspension of A. halophila cells was sonicated in 20 ml batches (sonication pulses 30 s, cooling intervals 2 min) for 1 min / 2 ml cell suspension. The suspension of E. halochloris cells was sonicated in 5 ml batches (sonication pulses 15 s; cooling intervals 2 min) for 1 min/1.5 ml cell suspension. The cell debris was removed by centrifugation at 28,000 g at 1°C for 30 min. The cell free extracts were stored at -75°C.

Methyltransferase activity assay

Reactions were carried out in 1.5 ml Eppendorf tubes with caps. The reaction mixture contained 25 µl of 0.1 M substrate (glycine, sarcosine or dimethylglycine), 25 µl of Buffer II (see above), 25 µl 4 mM S-adenosyl-L-methionine containing 45 nCi S-adenosyl-L-[methyl-^{14}C] methionine (Amersham) in 1/10 McIlvaine buffer (pH 3.0), and 25 µl enzyme sample (e.g. cell free extract). The reaction was initiated by adding the enzyme. The reaction mixture was incubated for 30 min. at 37°C and the reaction was stopped by adding 75 µl of charcoal suspension (133 g/l in 0.1 M acetic acid). The excess charcoal selectively adsorbs unreacted S-adenosyl methionine. The reaction mixtures were then incubated for 10 min at 0°C and centrifuged for 10 min in a Heraeus table top centrifuge. 75 µl of the supernatant was added to 4.5 ml of aqueous scintillant (Hionic-Fluor, Packard) and the radioactive methylation products were measured in a liquid scintillation counter (Beckman LS 6000 IC). The enzyme sample was diluted to keep the reaction on the linear range (radioactivity of the supernatant below 10,000 DPM).

The cell extracts typically contain the following activities.

Table 1. Methyltransferase activities of A. halophila and E. halochloris cell extracts on different substrates.

Organism	*Activity/DPM/30 min. incubation		
	Glycine	Sarcosine	Dimethylglycine
A. halophila	900	89,000	84,000
E. halochloris	19,000	76,000	854,000

* The activity is expressed in units based on the radioactivity of the reaction products after 30 min. incubation

Characterization of the methylation reaction products in the reaction mixtures

The reaction products were characterized by HPLC. The reaction mixture supernatants as described above were filtered after centrifugation through a Minisart NML 0.2 µm filter (Sartorius AG) and a 25 - 100 µl sample was analyzed on AminexHPX-87C cation exchange column (300 x 7.8 mm) (BioRad Laboratories). The HPLC system used was a Varian 500 equipped with a HP 1047(B) refractive index detector and a Waters VISP717 injector. A µbondapack C_{18}-precolumn was used in the system. 5 mM $CaSO_4$ was used as the eluent and the flow rate was 0.6 ml/min. A 1 mM mixture of sarcosine, dimethylglycine and betaine were used as standards. In order to detect the radioactive products formed in the enzymatic reaction, 200 µl fractions were collected during the chromatographic run. The fractions were analyzed in a liquid scintillation counter as described above. The dead-volume between the detector and sample outlet was determined by radioactive betaine and the retention time of the different fractions was calculated from the sample volume and eluent flow rate. The radioactivity of the fractions was plotted to the same Figure with the standards, which were used to identify the reaction products.

The results obtained by using A. halophila cell extracts are presented in Figure 1. The results indicate that all methylation products in the three-step reaction are formed during the incubation.

Example 2. Purification of glycine-sarcosine methyltransferase (GSMT) of E. halochloris

20 mM Tris-HCl buffer (pH 7.5) was used through the purification procedure unless otherwise stated. All the buffers used contained 1 mM dithiothreitol.

Step 1: Ammonium sulphate fractionation. 25 ml of cell free extract (as described in example 1) was diluted to 90 ml and saturated ammonium sulphate in 50 mM Tris-HCl, pH 7.5, was added to achieve 20 % saturation. The solution was incubated for 30 min at 0°C and centrifuged at 15,000 g. The precipitate was discarded and the supernatant purified further.

Step 2. Hydrophobic interaction chromatography. The supernatant from step 1 (105 ml) was applied to a Butyl Sepharose 4 FF (Pharmacia) (10 x 50 mm) column pre-equilibrated with 20 % (w/v) ammonium sulphate in 20 mM Tris-HCl, pH 7.5. The column was washed with 45 ml of 20% (w/v) ammonium sulphate in 20 mM Tris-HCl and eluted with a linear gradient of 20-0% ammonium sulphate. The volume of the gradient was 80 ml and the flow rate was 2 ml/min. Fractions of 3 ml were collected. The active fractions (40 ml) were pooled. The ammonium sulphate was removed by gel filtration (Sephadex G-25, Pharmacia).

Step 3. Ion exchange chromatography. The sample from step 2 (73 ml) was applied to a DEAE-Memsep 1000 HP (Millipore) (1.4 ml) column pre-equilibrated with 20 mM Tris-HCl, pH 7.5. The column was washed with 15 ml of buffer and eluted with a linear NaCl gradient (0 - 1 M). The volume of the gradient

was 60 ml and the flow rate was 3 ml/min. 2 ml fractions were collected The active fractions (8 ml) were pooled and concentrated by ultrafiltration (Amicon Centriplus 30; Ultrafree MC 10,000 NMWL filter unit Millipore) to 100 µl.

Step 4. Gel filtration. The concentrated sample from step 3 (100 µl) was applied to a Superose 12 HR 30 (Pharmacia) column 20 mM Tris-HCl, pH 7.5, containing 150 mM NaCl was used as the elution buffer with flow rate of 0.4 ml/min. 0.5 ml fractions were collected. The fractions containing glycine-sarcosine methyltransferase (GSMT) activity (1.5 ml) were collected and concentrated by ultrafiltration (Ultrafree MC 10,000 NWML filter unit, Millipore) to 100 µl.

The purity and the molecular weight were determined by gradient SDS-polyacrylamide gel electrophoresis according to the following procedure. Electrophoresis under denaturing conditions was carried out using pre-made polyacrylamide gel slabs (12 % Tris-glycine gel with 4 % stacking gel, Ready Gels, Biorad) according to the instructions of the manufacturer. Mid range molecular weight standard from Promega was used. Staining of the gel was performed with 0.25% (w/v) Coomassie Blue R-250 (Promega) in 50 % (v/v) methanol and 10% (v/v) acetic acid. Stained gels were destained with 10% (v/v) methanol and 5% acetic acid (Laemmli, 1970).

A typical SDS-gel of the purified E. halochloris GSMT is shown in Figure 2.

Example 3. Purification of sarcosine-dimethylglycine methyltransferase (SDMT) of Actinopolyspora halophila

An affinity purification method was developed for the purification of the sarcosine-dimethylglycine methyltransferase (SDMT). Purification of the protein was not achieved by standard purification methodology. The affinity column was prepared as follows. 5'AMP-Sepharose 4B

(Pharmacia) was treated with alkaline phosphatase to remove the phosphate group of the ligand. The gel was first swollen in water (5 ml of distilled water was used per 1 g of dry 5'AMP-Sepharose 4B). The swollen gel was then washed with 200 ml of distilled water. The gel was equilibrated with CIP-buffer (10 mM $MgCl_2$, 1 mM dithiothreitol, 50 mM NaCl, 10 mM Tris-HCl, pH 7.9). 10 µl (100 U) of CIP (Calf Intestinal Alkaline Phosphatase (Finnzymes)) was added to the gel. The gel was incubated for 2 h at 37°C with occasional shaking. The reaction was stopped by washing the gel with 20 mM Tris-HCl, pH 7.5.

18 ml of A. halophila cell free extract (as described in example 1) was applied to a column of adenosine-Sepharose (10 x 90 mm) pre-equilibrated with 20 mM Tris-HCl, pH 7.5. The column was washed with 20 mM Tris-HCl, pH 7.5, using 0.5 ml/min flow rate until the absorbence at 280 nm became constant. The protein bound to the column was eluted with 1 mM S-adenosyl methionine (20 mM Tris-HCl, pH 7.5). The active fractions were pooled (25.2 ml) and concentrated by ultrafiltration (Amicon Centriplus 30). The purity and the molecular weight was determined by gradient SDS-polyacrylamide gel electrophoresis as described above.

A typical SDS gel of the purified A. halophila SDMT is shown in Figure 2.

Example 4. Characterization of the properties of A. halophila SDMT protein

Determination of the molecular weight and sub-unit structure

Based on the SDS-electrophoresis (Figure 2), the molecular weight of A. halophila SDMT is approximately 32 kDa.

The molecular weight was also determined by gel filtration with Superose 12 HR 30 (Pharmacia) column. The flow rate was

0.4 ml/ml. The elution buffer was 20 mM Tris-HCl pH 7.5 containing 150 mM NaCl.

The molecular weight was calculated from a calibration curve made with a mixture of standard proteins. The mixture contained 0.5 mg/ml Blue Dextran (0.5 mg/ml), Ferritin (440 kDa), 7.0 mg/ml aldolase (158 kDa), 2.0 mg/ml ovalbumin (43 kDa) and 1.0 mg/ml chymotrypsinogen (25 kDa). The calculated molecular weight was 31.6 kDa, which indicates that the protein is a monomer.

The isoelectric point of the enzyme

The isoelectric focusing was performed with Pharmacia Phast system using gels with pH-gradient from pH 3 to 9 (IEF 3-9). A mixture of Pharmacia IEF standard proteins with pIs from 3.5 to 9.3 were used as standards. The gels were stained by silver staining as described in example 1. The results shown in Figure 3 show that the pI of the protein is approximately 4.1-4.2.

Substrate specificity

The activity of the purified protein was determined as described in example 1 with glycine, sarcosine and dimethyl glycine. The data presented in table 2 demonstrates that the isolated protein catalyzes step methylation reaction from sarcosine to dimethyl glycine and from dimethyl glycine to betaine.

Table 2. The activity of A. halophila SDMT on different substrates.

Substrate	*Activity (DPM/30 min.)
Glycine	0
Sarcosine	14,500
Dimethylglycine	45,800

* The activity is expressed in arbitrary units based on the radioactivity of the reaction products after 30 min. incubation.

pH Optimum

The pH-optimum of the two methylation reactions were determined by using following buffers: 0.1 M piperazine buffer, pH 5.0; 0.1 M Bis-Tris buffer, pH 6.0; 0.1 M Bis-Tris buffer, pH 7.0; 0.1 M Tris-HCl, pH 8.0; 0.1 M Tris-HCl, pH 9.0; 0.1 methanolamine, pH 10. The exact pH values of the reaction mixtures were measured. It can be concluded from Figure 4 that the pH optimum of the both enzyme reactions is at pH 7.5.

Temperature optimum

The temperature dependence of the enzymatic reactions were determined with sarcosine and dimethyl glycine. As seen in Figure 5. the temperature optimum is approximately 45 - 50°C. When the temperature is elevated above 50°C, the enzyme is rapidly inactivated.

Example 5 In vitro synthesis of betaine by using purified E. halochloris GSMT and A. halophila SDMT

The purified GSMT from E. halochloris and SDMT from A. halophila were concentrated by ultrafiltration (Ultrafree MC 10,000 NWML filter unit, Millipore) to protein concentrations 4.2 mg/ml and 5.6 mg/ml, respectively. The protein concentration was determined by measuring the absorbance at 280 nm and calculated by the formula: 1 mg protein/ml = 1.0 A_{280}.

The reaction mixture contained 50 µl 5.0 mM glycine, 50 µl 32 mM S-adenosyl methionine in water containing 640 nCi S-adenosyl-L-[methyl-^{14}C]-methionine (Amersham), 50 µl Buffer

II (see example 1), 25 µl GSMT of E. halochloris and 25 µl SDMT of A. halophila. The reaction was initiated by adding the enzymes. The reaction mixture was incubated for 2 h at 37°C and the reaction was stopped by adding 150 µl of charcoal suspension. The reaction mixtures were then incubated for 10 min at 0°C and centrifuged for 10 min in a Heraeus table top centrifuge. The supernatants were filtered through Minisart NML 0.2 µm filter (Sartorius AG). The identification of the reaction products was performed by HPLC as described in example 2.

The chromatogram is presented in Figure 6 and it shows peaks corresponding to the retention times of sarcosine, dimethylglycine and betaine.

Example 6. Isolation of the genes of E. halochloris GSMT and A. halophila SDMT

Determination of the N-terminal and internal peptide sequences

The N-terminal and tryptic peptides peptide sequences of the purified proteins were determined by using Perkin Elmer/Applied Biosystems Procise 494A protein sequencing system as described by Kerovuo et al., 1998. The peptide sequences obtained are shown in table 3.

Table 3. The peptide sequences obtained from purified E. halochloris GSMT and halophila SDMT. The sequences used to make the PCR primers are underlined.

Organism/SEQ ID NO:	Sequence
A. halophila	
SEQ ID NO:9	EKSYRTEDEFVDMYSNAVHTARDYYNSEDASNFYYHV (N-terminus)
SEQ ID NO:10	GSVLFTDPMASDDAK
SEQ ID NO:11	TGLRNYQAGN

SEQ ID NO:12	LXELGPILDRLHLDSG
SEQ ID NO:13	ELTRLGLQNIEFEDLSEYLPVHYGR
SEQ ID NO:14	VDISPETRILDLGSGYGA
E. halochloris.	
SEQ ID NO:15	NTTT/EEQDFGADPTKVRDTDAYTE
	(N-terminus)
SEQ ID NO:16	VRDTDHYTEEYVD
SEQ ID NO:17	DYTRRLMHEVGFQK
SEQ ID NO:18	ATYRDADPDFFLHVAEK
SEQ ID NO:19	VRDTDHYTEEYVDGFVDKWDDLID

Preparation and screening of the chromosomal gene banks

The genomic DNA from both microbes was isolated essentially as described in Ausubel et al. (1991). The chromosomal DNAs were partially digested with SacI and ligated to SacI digested dephosphorylated lambda ZapII arms (Stratagene, La Jolla, California, USA) and packaged to lambda particles using Gigapack III Gold packing extract (Stratagene, La Jolla California, USA) according to protocol provided by manufacturer. The chromosomal DNA isolated from the organisms was used as the template DNA in the PCR reactions.

The probes were made by PCR using following degenerate primers. The primers were designed according to Sambrook et al. (1989).

A. halophila
SEQ ID NO:20
5'-GA(A/G)GA(C/T)GA(A/G)TT(C/T)GTIGA(C/T)ATG T-3'

SEQ ID NO:21
(5'-(C/T)TG(A/G)TT(T/A/G)AT(T/C)TC(G/A)AA(T/C)TC(A/G)TC-3')

E. halochloris
SEQ ID NO:22
5'-GA(A/G)CA(A/G)GA(T/C)TT(T/C)GGIGCIGA(T/C) CC-3'

SEQ ID NO:23
5'-A(A/G)(A/G)AA(A/G)AA(A/G)TCIGG(A/G)TCIGC(A/G)TC-3'

The amplification was performed under the following conditions.

A. halophila- 3 cycles of 1 min. at 94°C for denaturation, 1 min at 37°C annealing and 2 min at 72°C for synthesis, 32 cycles of 1 min at 94° C for denaturation, 1 min at 46°C annealing and 2 min at 72°C for synthesis. AmpliTac DNA Polymerase (Perkin Elmer) was used in the reaction. 1 mM $MgCl_2$ was added to the reaction mixture. Otherwise standard reagents were used.

E. halochloris - 34 cycles of 1 min at 94° for denaturation, 1 min at 42°C annealing and 2 min at 72°C for synthesis. AmpliTac DNA Polymerase (Perkin Elmer) was used in the reaction. 1 % (v/v) formamide, 1% (v/v) dimethyl sulfoxide and 6 mM $MgCl_2$ were added to the reaction mixture. Gitsch-buffer (reference) was used in the reaction.

The PCR-fragments obtained were labeled with rediprime DNA labelling system (Amersham Life Science) according to the instructions given by the manufacturer.

A total of 50,000 plaques of both libraries were screened and the positive lambda clones were cored and excised with ExAssist helper phage (Stratagene, San Diego, USA) to obtain phagemids. CsCl-gradient purified (Sambrook et al., 1989) plasmid DNAs were used in DNA sequencing (Zagursky et al., 1986). The lengths of the cloned fragments were 3.5 kb (A. halophila) and 5.0 kb (E. halochloris).

Analysis of the sequencing data

E. halochloris - Sequence analysis of the DNA fragments indicated that the E. halochloris clone contains 3 ORFs. On

the basis of the peptide sequences it can be concluded that the E. halochloris GSMT is encoded by the first ORF of the fragment. In addition, based on the sequence homology with the A. halophila SDMT, it can be concluded that the second ORF of E. halochloris clone contains a SDMT gene. This has also been demonstrated by expressing the gene in E. coli (Example 8).

A. halophila - Sequence analysis of the DNA fragments indicated that the A. halophila clone contains 2 ORFs. A. halophila SDMT protein is coded by the 3'-end of the first ORF of the A. halophila clone. The 5'-end of the same ORF is very homologous to the E. halochloris GSMT. These data indicate that the GSMT and SDMT are transcribed from a single gene and the protein isolated from the organism is a processing product.

In addition to the methyltransferases, the "betaine operon" codes for a third gene which is homologous to number of S-adenosyl-methionine synthases. The operon structure is schematically shown in Figure 7. The nucleotide and amino acids sequences of the cloned genes have been shown in Figures 8 and 9.

Example 7 Expression of E. halochloris GSMT in E. coli

Expression of the gene

The gene coding for the E. halochloris GSMT was amplified by PCR. The purified plasmid used for DNA sequencing in example 6 was used as the template for the PCR reaction. The following primers were used in the PCR reaction:

primer 1:
5'-CGGACCATGGATACGACTACTGAGCAG-3' (SEQ ID NO:24)
(5'-end oligonucleotide) and

primer 2:
5'-GCTCAGATCTGTCCTCCTCCCGATATTCCTTCTC-3' (SEQ ID NO:25)
(3'-end oligonucleotide)

The 3'-end of the primers are homologous to the 5'- and 3'-end of the GSMT gene. The 5'-end oligonucleotide hybridizes to position 221-241 and the 3'-end to the position 1001-1024. (See Figure 8). The primer hybridizing to the 5'-end contains an extra NcoI restriction site such that the nucleotide A at position 224 in Figure 8 is replaced by the nucleotide G in the primer and the 3'-end primer contains a BglII site which were used for cloning.

The amplification was performed in the following conditions: 34 cycles of 1 min at 94°C for denaturation. 1 min at 50 °C annealing and 2 min at 72°C for synthesis. Pfu Polymerase (Stratagene) was used in the reaction. The amplified fragments were purified with Qiaquick DNA purification Kit (Qiagen, Santa Clara, USA) and ligated into NcoI/BglII cut PQE-60 expression vectors (Qiagen, Santa Clara, USA). A schematic presentation of the plasmid (pEGSM) is shown in Figure 10. Competent XL-1 Blue MPF' cells were transformed with this ligation mix according to Hanahan et al. (1983). Plasmid minipreps of the transformants were prepared and the presence of the insert was established by cutting the plasmids with Nco I and Bgl II and by separating the resulting fragments by agarose gel electrophoresis.

GSMT transformants were grown overnight in 2.5 ml of LB broth containing 100 µg/ml ampicillin. As a control, E. coli XLI Blue MRF' transformed with the PQE-60 without the insert was grown.

0.5 ml was inoculated to 1.5 ml of LB broth with ampicillin. The cultures were grown at 200 rpm for 30 min at 37°C and isopropyl-ß-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM to induce the enzyme synthesis. After 3

h 30 min the cells were separated by centrifugation (1,000 g, 3 min).

The cell pellet was suspended in 100 µl Buffer II containing 1 mM PMSF (See example 1) and the cells were disrupted with a MSE Soniprep 150 sonicator. The cell suspension was sonicated with sonication pulses of 5 s for 10 s. The samples were cooled on ice between the pulses. The cell debris was removed by centrifugation at 13,000 rpm for 30 min at 4°C in a Heraeus table top centrifuge. The activities of the supernatants were determined as in example 1.

The activity of the cell extracts of the transformants was assayed as described in example 1. The activities using glycine and sarcosine as substrates were typically 3,000-5,000 dpm/30 min and 1,000-2,000 dpm/30 min., respectively.

Salt tolerance of E. coli clones expressing the E. halochloris GSMT

The strains used in these tests were the positive clone designated EGSM and E. coli XLI Blue MRF' transformed with the cloning vector PQE-60. The growth medium used in this test was the synthetic medium MM63 described by Larsen et al. (1987) supplemented with 1.5 mil/l of vitamin solution VA (Imhoff and Trüper, 1977) and 100 µl/ml ampicillin.

The bacterial strains were grown to mid-exponential growth phases with shaking at 180 rpm at 37°C and centrifuged (1,000 g, 15 min). The cells were resuspended in the growth medium to absorbance of 0.9 at 600 nm.

The automatic turbidimetric system Bioscreen (Labsystems) was used for monitoring the growth. 30 µl was inoculated to 300 µl of the medium with 0 or 0.2 M NaCl and 1 mM IPTG (final concentrations). Cells were grown at 37°C with intensive and

continuous shaking. Growth was followed by absorbance measurements in Bioscreen at 600 nm.

The growth curves of EGSM and E. coli transformed with PQE-60 in a medium without added NaCl are presented in and with 0.2 M NaCl in Figure 11. It can be concluded from the data that EGSM has an increased tolerance towards osmotic stress.

Example 8. Expression of E. halochloris SDMT in E. coli

The gene encoding the E. halochloris SDMT was amplified by PCR. The purified plasmid used for DNA sequencing in example 6 was used as the template for the PCR reaction. The following primers were used in the PCR reaction:

primer 3:
5'-GCATGCCATGGCGACGCGCTACGACGATCAA-3' (SEQ ID NO:26)
(5 -end oligonucleotide) and

primer 4:
5'-GGGAAGATCTCCCTTTGCGGAAGTAAAAGATACC-3' (SEQ ID NO:27)
(3'-end oligonucleotide)

The primers are homologous to the 5'- and 3'-end of the E. halochloris SDMT gene. The 5'-end oligonucleotide hybridizes to position 1031-1054 and the 3'-end to the position 1844-1867 (Figure 8). The primer hybridizing to the 5'-end contains an extra NcoI restriction site and the 3'-end primer a BglII site which were used for cloning of the fragment.

The preparation of the DNA construct and the cultivation of the positive E. coli clones were done according to example 7. A schematic presentation of the expression plasmid (pESDM) is shown in Figure 10.

The cultivations and preparation of the cell-free extracts were performed essentially as described in example 7. The

sonication of the cell extract pulses was shortened to 3 x 2 second intervals (total sonication time was 6 s).

The activity of the cell extracts of the transformants was assayed as described in example 1. The activities using sarcosine and dimethyl glycine as substrates were typically 20,000 dpm/30 min. with both substrates.

Example 9. Co-expression of E. halochloris GSMT and SDMT in E. coli

The DNA construct made for this experiment contains both GSMT and SDMT genes separated by a short (3 nucleotides long) linker. The DNA fragment was obtained by amplification of the purified plasmid used for DNA sequencing in example 6.

The following primers were used in the amplification:

primer 1:
5'-CGGACCATGGATACGACTACTGAGCAG-3' (SEQ ID NO:24)
(5'-end oligonucleotide) and

primer 4:
5'-GGGAAGATCTCCCTTTGCGGAAGTAAAAGATACC-3' (SEQ ID NO:27)
(3'-end oligonucleotide)

The primers are homologous to the 5'-end of the E. halochloris GSMT and the 3'-end of the E. halochloris SDMT gene. The 5'-end oligonucleotide hybridizes to position 221-241 and the 3'-end to the position 1844-1867 (See Figure 8). The primer hybridizing to the 5'-end contains an extra NcoI restriction site and the 3'-end primer a BglII site.

The preparation of the DNA construct and the cultivation of the positive E. coli clones were done according to example 7. A schematic presentation of the expression plasmid (pEhFU) is shown in Figure 10. The induction and preparation of the

cell-free extracts was performed essentially as described in example 9

The enzymatic activities were assayed as described in example I . The cell extracts of the transformants clearly showed activity with glycine, sarcosine and dimethylglycine. The activities of the cell extracts with the three substrates were all over 20,000 dpm/30 min.

Salt tolerance E. coli clones expressing the E. halochloris GSMT and SDMT

The test was performed essentially as described in example 7. The positive clone designated EhFU was used in this test.

The growth curves of EhFU and E. coli with PQE-60 in a medium without added NaCl are presented in and with 0.2 M NaCl in Figure 12. It can be concluded from the data that EhFU has an increased tolerance to osmotic stress.

Betaine synthesis in E. coli clones expressing the E. halochloris GSMT and SDMT

The growth medium used in this test was the synthetic medium MM63 described by Larsen et al. (1987) supplemented with 1.5 ml/l of vitamin solution VA (Imboff and Trüper, 1977) and 100 µl/ml ampicillin. The medium contained 1% (w/v) glucose.

The clone EhFU (Figure 10) and the control strain (E coli XL1-Blue MRF' transformed with the cloning vector PQE-60) were grown to mid-exponential phase with shaking at 180 rpm at 37°C. The cells were centrifuged at 1,000 g for 10 min and resuspended in the growth medium so that the turbidity A_{600t} was 0.640. 5 ml of this cell suspension was inoculated to 50 ml of media containing 0.22 or 0.33 M NaCl and 25 mM L-methionine. The bacterial strains were grown for 2 h with shaking at 180 rpm at 37°C and 1 mM IPTG was added. Growth was followed by measuring the turbidity at 600 nm. The cells

were grown to early stationary phase. Cells from 45 ml of culture were harvested by centrifugation at 1,000 g for 15 min and washed once with the growth medium without glucose.

The cell pellets were suspended in 2 ml of water and kept in a boiling water bath for 10 min. The suspension was centrifuged for 15 min at 23,000 g and the supernatant collected and the pellet resuspended in water. This extraction was repeated twice. The three supernatants were combined. The volumes of supernatants were measured and the supernatants were filtered and analyzed by HPLC as described in example 1.

A similar experiment was performed also without added L-methionine.

The betaine produced inside the cells is presented in table 4.

Table 4. The amount of betaine synthesized inside the cells in 1 ml of culture when grown in MM63

		Betaine concentration in the cells ($\mu g/A_{600}ml$)	
Transformant	NaCl (mol/l)	Medium supplemented with 25 MM methionine	No methionine added
EhFU	0	n.d.	1.0
	0.2	3.5	1.1
	0.3	2.4	0.9
PQE-60	0.2	0	0
	0.3	0	0

The results show that the E. coli clone expressing the E. halochloris GSMT and SDMT genes synthesizes betaine in these cells. The highest amount of betaine synthesized corresponds

roughly to 1% (0.2 M NaCl) and 0.5 % (0.3 M NaCl) of the cell dry weight

Example 10. Expression of the DNA fragment encoding the protein isolated as A. halophila SDMT in E. coli

The gene sequencing results revealed that a single gene codes for the A halophila GSMT and SDMT. The fusion protein was not, however, successfully purified from the A. halophila cell extracts. Instead a protein with SDMT activity was isolated. In this experiment the corresponding part of the GSMT-SDMT gene is expressed in E. coli.

The gene fragment encoding the SDMT enzyme activity was amplified by PCR. The genomic DNA from A. halophila isolated in example 6 was used as the template for the PCR reaction. The following primers were used in the amplification:

primer 5:
5'-GCTGCCATGGAGAAGAGCTACCGCACCGAG-3' (SEQ ID NO:28)
(5'-end oligonucleotide) and

primer 6:
5'-GGGAAGATCTTGCCCTGGCGTGGATGATGCCCCA-3' (SEQ ID NO:29)
(3-end oligonucleotide)

The primers are homologous to the 5'- and 3'-end of the ASDMT gene. The 5'-end oligonucleotide hybridizes to position 1048-1068 and the 3'-end to the position 1879-1902 (See Figure 9). The primer hybridizing to the 5'-end contains an extra NcoI restriction site and the 3'-end primer a BglII site.

The preparation of the DNA construct and the cultivation of the positive E. coli clones were done according to example 7. A schematic presentation of the expression plasmid (pASDM) is shown in Figure 10. The induction and preparation of the cell free extracts was performed essentially as described in example 9.

The enzymatic activities were assayed as described in example 1. The cell extracts of the transformants clearly showed SDMT activity. The activities on sarcosine and dimethyl glycine were typically 20,000 dpm/30 min. with both substrates. There was no activity on glycine.

Example 11. Expression of A. halophila GSMT-SDMT fusion protein in E. coli

The gene fragment encoding GSMT and SDMT of A. halophila was amplified by PCR. Primers used were:

primer 7:
5'-CATGCCATGGCCAAGAGCGTGGACGATCTT-3' (SEQ ID NO:30)
(5'-end oligonucleotide) and

primer 6:
5'-GGGAAGATCTTGCCCTGGCGTGGATGATGCCCCA-3' (SEQ ID NO:29)
(3-end oligonucleotide)

The primers are homologous to the 5'- and 3'-end of the AGSMT-ASDMT gene. The 5'-end oligonucleotide hybridizes to position 208-231 and the 3'-end to the position 1879-1902 (See Figure 9). The primer hybridizing to the 5'-end contains an extra NcoI restriction site such that the nucleotide A at position 211 in Figure 9 is replaced by the nucleotide G in the primer and the 3'-end primer contains a BglII site.

The purified plasmid used for DNA-sequencing in example 6 was used as a template for the PCR reaction. The amplification was performed in following conditions: 34 cycles of 30 s at 94°C for denaturation, 1 min at 50°C annealing and 2 min at 72°C for synthesis. Ligation of the amplification product into NcoI/BglII cut PQE-60 and the transformation of XL-1 Blue MRF' cells was performed as in example 7. A schematic presentation of the expression plasmid is shown in Figure 10.

The induction and preparation of the cell-free extracts was performed essentially as in example 9 except that the sonication pulses were shortened to 2 s and the total sonication time to 6 s. The cell-free extract was analyzed by SDS-polyacrylamide gel electrophoresis as in example 2. The pellet from the centrifuged suspension was suspended to 10 mM Tris-HCl-buffer, pH 8.0 containing 8 M urea and 0.1 M Na_3PO_4 to solubilize the proteins of the pellet and centrifuged for 15 min in a Heraeus table top centrifuge at 13,000 rpm. The supernatants were analyzed by SDS-polyacrylamide gel electrophoresis as in example 2.

The enzymatic activities were assayed as described in example 1. The cell extracts of the transformants clearly showed SDMT activity. The activities on sarcosine and dimethyl glycine were typically 10,000 dpm/30 min and 20,000 dpm/30 min., respectively. There was no activity on glycine.

The SDS-polyacrylamide gel of the cell-free extract showed no major protein band of correct size. However, the insoluble pellet solubilized with 8 M urea showed a major band corresponding to the molecular weight of the GSMT-SDMT fusion protein. The results indicate that when A. halophila GSMT-SDMT is over-expressed in E. coli it forms inclusion bodies. However, a fraction of the protein - which corresponds the SDMT - is proteolytically cleaved and remains soluble in the cells.

Example 12. Expression of *E. halochloris* GSMT and SDMT in tobacco and potato

Tobacco and potato plants can be transformed by *Agrobacterium* mediated transformation system. Identical DNA construct can be used for both plants.

The GSMT gene is first transformed into the plant using a plasmid containing a kanamycin resistance marker. Positive transgenic plants obtained by screening for the enzyme

activity are then used as host plants for second transformation of the SDMT gene. Another selection marker, hygromycin selection is used in the second transformation. Experiments are performed using stable transformants of the F_1 generation.

The genes of *E. halochloris* GSMT and SDMT are amplified by PCR by using plasmid pEFU (see example 10) as the template. The primers used hybridize to the same regions of the DNA as shown in Fig. 8 (GSMT: primer 1 and primer 2; SDMT; primer 3 and primer 4). The final DNA constructs are made using suitable restriction sites to transfer the genes to plant transformation vectors. PBin 19 based pGPTV vectors (Becker et al, 1992) are used which have a strong 35S promoter and the CaMv polyadenylation signal.

The resulting plasmids are transformed to *A. tumefaciens* strains. Strain EHA 105 (Hood, E.E. et al., (1993) is used as a vector to transform tobacco basically as described by Rogers et al. (1986). Strain C1C58p-GV3850 (Zambryski et al., (1983); Van Larabece et al., (1974)) is used as an alternative host to transform potato *Solanum tuberosum* (Desiré) according to Dietze et al. (1995).

The transformants are analyzed by Southern blot analysis to check for the presence of the genes. PCR-amplified, DIG-labelled (Boehringer) 200 bp gene fragments are used as a probe. The enzymatic activities of the cell extracts of transgenic plants and the levels of sarcosine, dimethyl glycine and betaine are analyzed as described in Example 1.

Stress tolerance, for example, tolerance to drought, salinity, cold or freezing, resistance to pathogens, etc., is determined according to methods known in the art, for example, methods described in the technological background section of the present application.

Example 13. Expression of *E. halochloris* GSMT and SDMT in rice

The plasmid constructions described in Example 13 are also used to transfer the GSMT and SDMT genes to rice by particle bombardment. The GSMT are transferred to rice first and positive regenerated transformants are used as host plants for the SDMT transformation.

The following procedure are used. Immature *Oryza sativa* embryos of the Japonica variety Taipei 309 are aseptically isolated 10-14 days after pollination from greenhouse plants and plated scutulum site up on solid MS medium (Murashige and Skoog, 1962) containing 3% sucrose, 2 mg/l 2,4-dichlorophenoxyacetic acid and 50 mg/l cefotaxime (MS1). After 4-6 days (28°C, darkness) embryos are transferred to solid MS medium containing 10% sucrose, 2 mg/l 2,4-dichlorophenoxyacetic acid and 50 mg/l cefotaxime (MS2) and subjected within 1 hour to microprojectile bombardment with a particle inflow gun. The DNA fragment containing the mehtyltransferase gene and the selective marker (5 μg) is precipitated on 1-3 mm gold particles (Aldrich) as described by Vain et al., (1993). Gold particles (400 mg per bombardment) are accelerated to the target with a particle inflow gun (Finer et al., (1992) at a pressure of 6 bar. Embryos are placed 16 cm below the syringe filter. Twenty four hours post-bombardment embryos are subjected to selection on solid media (containing hygromycin or kanamycin) and incubated at 28°C in the dark.

After one week embryos are transferred to a liquid selection media, R2 medium (Ohira et al., (1973) supplemented with: 3% sucrose, 1 mg/l thiamine, 1 mg/l 2,4-dichlorophenoxyacetic acid, 50 mg/l cefotaxime and 20 mg/l hygromycin B or kanamycin. The embryos are incubated with shaking at 28°C in the dark and subcultured weekly. Developing calli are isolated 3 to 6 weeks later, and transferred to a callus increasing media (R2 medium supplemented with: 6% sucrose, MS

vitamins, 100 mg/1 inositol, 2 mg/1 2,4-dichlorophenoxyacetic acid, 50 mg/1 cefotaxime and 20 mg/1 hygromycin B kanamycin). The calli are incubated in this media at 28°C in the dark and subcultured weekly.

Resistant calli are transferred to solid R2 regeneration media supplemented with 2% sucrose, 3% sorbitol, 20 mg/1 hygromycin B, 1 mg/1 zeatin, 0.5 mg/1 indole-3-acetic acid, MS vitamins and 0.65% agarose. The callus tissue is maintained at 28°C with 12 h of light in order to enhance shoot formation. The calli are then subcultured every 3 weeks until shoots had reached a length of 2-3 cm. They are transferred to half-strength MS rooting medium without hormones, supplemented with 1.5% sucrose and 0.3% gelriteR (Sigma). After 2-4 weeks of cultivation, plantlets are transferred directly to the green-house and planted in soil. Plantlets are grown in 7 liter aquaculture pots with fertilizer enriched earth, 3 plants per pot (day: 12 h, 28°C, 80% humidity; night: 12h, 21°X, 60% humidity) until they flower and set seeds.

To check the presence of the transgene, complexity of insertion(s) and number of copies present, Southern blot analysis is performed as described previously (Burkhardt et al., 1997). A PCR amplified, DIG-labelled (Boehringer) 200-bp fragment of the coding region of the GSMT or SDMT genes is used as a probe.

The enzymatic activities of the cell extracts of transgenic plants and the levels of sarcosine, dimethyl glycine and betaine are analyzed as described in Example 1.

Stress tolerance, for example, tolerance to drought, salinity, cold or freezing, resistance to pathogens, etc., is determined according to methods known in the art, for example, methods described in the technological background section of the present application.

Example 14. Expression of the *E. halochloris* GSMT and SDMT in yeast

pYX242 plasmid (R&D systems, USA) was used for expressing the GSMT and SDMT genes in *Saccharomyces cerevisiae*. The plasmid used (pYX242) is a *E. coli-Saccharomyces cerevisiae* shuttle vector containing a bacterial origin of replication and ampicillin resistance gene, a yeast (*S. cerevisiae*) origin of replication from 2μm DNA, and the yeast LEU2 gene for selection in yeast. The two genes are expressed under the yeast triose phosphate isomerase (TPI) promoter.

The DNA of the plasmid pEFU described in Example 10 was used as the template of PCR reactions. The primers used hybridize to DNA sequences shown in Fig. 8 (primer 1 and primer 4) and thus amplify both the GSMT and SDMT genes. The PCR fragment was ligated to the promoter of the expression plasmid with standard methods. A fragment containing a TPI transcription terminator and a fragment containing the TPI promoter was ligated between the two genes. Thus, both genes are expressed under the TPI promoter. The primers used in the amplification of the fragments ligated in the expression plasmid contained suitable restriction sites that were used in the cloning. *S. cerevisiae* GRF18 (MATα, leu2-3, 11, his3-11.15) was used as the host for transformation. The transformation was performed according to Ito et al. using the standard lithium chloride procedure (Ito et al., (1983) Bacteriol. 153, 163-170) using the LEU2 marker of pYX242 for selection.

The transformants were grown in YNB-medium supplemented with amino acid mixture without leucine (R & D systems product manual). The cultivation was done overnight at 30°C by shaking at 180 rpm. 5 ml of culture supernatant was centrifuged (1,700 g, 10 min) and the cell pellet was suspended in 200 μl of the assay buffer (see example 1) supplemented with 1 mM PMSF. The cells were broken by vortexing with glass beads (10 x 1 min intervals). The cells

were kept on ice between the pulses. The cell debris was centrifuged down (30 min, 4°C, 10,000g).

The methylase activities were assayed from the supernatant as described in example 1. The relative enzyme activities on different substrates were the following: glycine - 17,000 dpm/30 min; sarcosine - 71,000 dpm/30 min and dimethyl glycine - 530,000 dpm/30 min.

References

Agboma, P. (1977) PhD Thesis, University of Helsinki, Department of Plant production, Section of Crop Husbandry

Augustine, P. and McNaughton, J.L. (1966) Proceedings of the Maryland Nutrition Conference, March 21-22, pp. 31-36

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G. Smith, J.A. and Struhl, K. (1991) Current Protocols in Molecular Biology. New York:Wiley

Barak, A.J., Beckenhauer, H.C., Junnila, M. Turna, D.J. (1993) Alcoholism: Clinical Exp. Res. 17, 552-555

Becker, D., Kemper, E. Schell, J., Masterson, R., (1992) Plant Molecular Biology 20, 1195-1197

Bhandari, P., Gowrishankar, J., (1997) J. Bacteriol. 179, 4403-4406

Blackwell, J.R., Horgan R. (1991) FEBS Letters 295, 10-12

Blunden, G., Wilson, S., Romero, M.A., Wu, Y., Jenkins, T., Tyihar, E. (1996) Chem. Listy 90, 679-706

Boch et al. (1996) J. Bacteriol. 178, 5121-5129

Boch, J., Kempf, B. Bremer, E. (1994) J. bacteriol. 176, 5364-5371

Bohnert, H.J., Jensen, R.G. (1996) Tibtech 14, 89-97

Brouquisse, R., Weigel, P., Rhodes, D., Yocum, C.F., Hanson A.D. (1989) Plant Physiol. 90, 322-329

Burkhardt, P.K., Beyer, P., Wuenn, J., Kloeti, A., Armstrong, G.A., Schledz, M., von Lintig, J., Potrykus, I. (1997) Plant Journal 11, 1071-1078

Chambers, S. and Kunin, C.M (1985) J. Infectious Diseases 152, 1308-1316

Coughlan, S., Heber, U. (1982) Planta 156, 62-69

Csonka L.N. (1989) Microbiological reviews 53, 121-147

Deshnium, P., Gombos, Z., Nishiyama, Y., Murata, N. (1997) J. Bacteriol. 179, 339-344

Dietze J., Blau, A. and Willmitzer, L. (1995) in Gene transfer in plants, pp.24-29, Potrykus, I., and Spangenberg, G. (eds), Springer-Verlag, Berlin

Ferket, P.R. (1994) Proceedings of the Smith Beecham Animal Health Poresetter Seminar, January 10, pp. 1-7

Finer et al. (1992) Plant Cell Rep., 11, 323-328

Galinski, E.A. (1993) Experentia 49, 487-496

Galinski, E.A. and Trüper, H.G. (1982) FEMS Microbiol. Lett. 13:357-360

Galinski, E.A., Trüper, H.G. (1994) FEMS Mircrobiol. Rev. 15, 95-108

Glaasker E., Konongs, W., E., Poolman B (1996) J. Bacteriol. 178, 575-582

Gorham, J. (1995) Betaines in higher plants - biosynthesis and role in stress metabolism. In Amino acids and their derivatives in higher plants, ed. R.M. Wallsgrove, pp. 171-203. Cambridge: Cambridge University Press

Guo, Y., Zhang, L., Xiao, G., Cao, S., Gu, D., Tian, W., Chen. S. (1997) Science in China 40, 496-501

Hanahan, D. (1983) Studies on transformaiton of Escherichia coli with plasmids. J. Mol., Biol. 166:557-580

Hanson, A.D., May, A.M., Grumet, R., bode, J., Jamison, G.C. (1985) Proc. Natl., Acad. Sci. USA 82, 3678-3682

Hayashi, H., Alia, Mustardy, L., Deshnium, P., Ida, M., Murata, M. (1997) Plant. J. 12, 133-142

Holmberg, N. (1996) PhD Thesis, Department of Pure and Applied Biochemistry, Lund University

Holmström, K-O, Welin, B., Mandal, A., Kristiansdottir, I., Teeri, Y.Y., Lamark, T., Ström, A., Palva, T.E. (1994) The Plant J. 6, 749-758

Hood, E.E., Gelvin, S.B., Melchers, L.S. Hoekama, A. (1993) Transgenic Research 2, 208-218

Imhoff, J.F. & Trüper, H.G. (1997) Arch. Microbiol. 114:115-121

Ishitani, M., Arakawa, K., Mizuno, K., Kishitani, S., Takabe, T (1993) Plant Cell Physiol., 34, 493-495

Ito et al. (1983) Bacteriol. 153, 163-170

JP patent 8-103267

JP patent 8-266179

JP patent 87,190,078 (Kojima, Y., Aisui, S., Ando, M. (1987)

Kapper, R., Kempf, B., Bremer, E., (1996) J. Bacteriol. 178, 5071-5079

Kempf, V., Bremer, E. (1995) J. Biol. Chem. 270, 16701-16713

Kishitani, S., Watanabe, K., Yasuda, S. Arakawa, K., Takabe, T. (1994) Plant, Cell and Environm. 17, 89-95

Kerovuo, J., Lauraeus, M., Nurminen, P., Kalkkinen, N, Apajalahti, J. (1998) Appl. Environm. Microbiol. 64, 2079-2085

Kuipers, O.P., de Ruyter G.G.A., Kleerebezem, M., de Vos, W. (1997) Tibtech 15, 135-140

Laemmli, U.K. (1970) Nature 227:680-685

Lamark, T., Kaasen I., Eshoo, M.W., Falkenberg, P., McDoygall, J., Ström, A.R. (1991) Mol. Microbiol. 5, 1049-1064

Larsen, P.I., Sydne, L.K., Landfald, B. & Ström, A.R. (1987) Arch. Microbiol. 147:1-7

Lilius, G., Holmberg, N., Bulow, L. (1996) Biotechnology 14, 177-180

Markham, G.D., DeParasis, J and Gatmaintan, J. (1984) J. Biol. Chem. 259, 14505-14507

McGue, K.F., Hanson A.D. (1990) Tibtech 8, 358-363

McGue, K., Hanson, A. (1992) Plant Mol. Biol. 18, 1-11

Murashige, T. and Skoog, F. (1962) Physiol. Plant 15, 473-497

Mäkelä, P., Mantila, J., Hinkkanen R., Pehu, E., Peltonen-Sainio, P. (1996) J. Agronomy and Crop Science 176, 223-234

Nagasawa, T., Mori, N., Tani, Y., Ogata, K. (1976) Agr. Biol. Chem. 40, 1743-1749

Nakamura, T., Yokota, S., Moramoto, Y., Tsutsui, K., Ogura, Y., Fukui, K., Takabe, T. (1997) Plant J. 11, 1115-1120

Nomura, M., Ishitani, M., Takabe, T., Rai, A.K., Takabe, T. (1995) Plant Physiol. 107, 703-708

Nomura, M., Muramoto, Y., Yasuda, S., Takabe, T., Kishitani, S. (1995) Uephytica 83, 247-250

Ohira et al. (1973) Plant Cellk Physiol. 14, 113-121

Papagiorgiou, G.C., Murata, N. (1995) Photosynt. Res. 44, 243-252

Perroud, B., Rudulier, D. (1985) J. Bacteriol. 161, 393-401

Peter, H., Burkovski, A., Krämer, R. (1996) J. Bacteriol. 178, 5229-5234

Pocard, J-A., Vincent, N., Boncompagni, E., Tombras Smith L., Poggi, M-C., LeRudulier, D. (1997) Microbiology 143, 1369-1379

Rathinasabapathi, B. et al., (1997) Proc. Natl. Acad. Sci. USA 94, 3454-3458

Rhodes, D. and Hanson, A.D. (1993) Ann. Rev. Plant Physiol. Plant Mol. Biol. 44, 357-384

Roberts, M.F., Lai, M-C., Gunsalus, R.P. (1992) J. Bacteriol. 174, 6688-6693

Rogers S.G. and Fraley, R.T. (1986) Methods in Enzymol 118, 627-684

Rozwadovski, K., Kchachatourians, G.G., Slvaraj, G. (1991) J. Bacteriol. 173, 472-478

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) Molecular cloning: A Laboratory Manual. Cold Spring Harbor, NY, USA

Sarvas, M. (1994) in Gene Expression in recombinant Bacillus In Gene Expression in Recombinant Plants, pp. 53-120. Ed. A. Smith, Marcel Dekker, New York

Saunderson, C.L. and MacKinlay, J. (1990) Brit. J. Nutr. 63, 339-349

Sehgal, S.N. and Gibbons, N.E. (1960) Can. J. Microbiol. 10:165-169

Severin, J., Wohlfarth, A., Galinski, E.A. (1992) J. Gen. Microbiol. 138, 1629-1638

Smirnova, G. and Oktyabrsky, O. (1995) Arch. Microbiol 163, 76-78

Tani, Y., Mori, N., Ogata, K., Yamada, H. (1979) Agrig. Biol. Chem. 43, 815-820

Tschichholz, I. and Trüper, H.G., (1990) FEMS Microbiol. Ecol. 73:181-186

Tschichholz-Mikus, I. (1994) PhD Thesis, Rheinischen Friedrich-Weilhelms.Universität, Bonn

US Patent 4,245,050 (Nakanisch t., Machida, Y. (1981))

Vain et al. (1993) Plant Cell Rep. 12, 84-88

Van Larabece, N., Engler, G. Holsters, M., Van den Elsacer, S., Zaenen, I., Schilperoot, R.A., Schell, J. (1974)Nature 252, 169-170

Weretilnyk, E., Hanson A. (1990) Proc. Natl., Acad. Sci USA 87,2745-2749.

Weretilnyk, E. Hanson, A. (1989) Arch. Biochem. Biophys. 271, 56-63

Wu, Y., Jenkins, T., Blunden, G., Whapam, C., Hankins, S.D. (1997) Fundam. Appl. Nematol. 20, 99-102

Wyn Jones, R.g., Stirey, r., Leigh, R.A., Ahmad, N., Pollard, A. (1997) A hypothesis on cytoplasmic osmoregulation. In Regulation of cell membrane activities in higher plants, Eds. E. Marre and O. Ciferri, pp. 121-136. Amsterdam: Elsevier/North Holland

Xin, W.B., Rajashekar, C.B., Zhou, H. (1996) Plant physiol. VIII.N2,P230

Yancey, P.H., Clark, M.E., Hand, S.C., Bowlus, R.d., Somero, G.N. (1982) Science 217, 1214-1222

Yocum, R.R, Perkins, J.B., Howitt, C.L. and Pero, J. (1996) J. Bacteriol. 178, 4604-4610

Zagursky, R.J., Berman, L.M., Baumeister, K., Lomax, N. (1986) Gene Anal. Techn. 2, 89-94

Zambryski, P., Joons, H., Genetello, C., Leemans, J. Van Montagu, M., Schell, J. (1983) EMBO J. 12, 2143-2150

What is claim d:

1. A nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c).

2. A nucleic acid molecule according to claim 1 which comprises the DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1.

3. A nucleic acid molecule according to claim 1 which comprises the DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5.

4. A nucleic acid molecule according to claim 1 which comprises the DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1.

5. A nucleic acid molecule according to claim 1 which comprises the DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5.

6. A nucleic acid molecule according to claim 1 which comprises the DNA sequence from nucleotide 208 to 1902 of SEQ ID NO:1.

7. A nucleic acid molecule according to claim 1 which comprises the DNA sequence from nucleotide 221 to 1867 of SEQ ID NO:5.

8. A nucleic acid molecule according to claim 1 which comprises the DNA sequence from nucleotide 208 to 2722 of SEQ ID NO:1.

9. A nucleic acid molecule according to claim 1 which comprises the DNA sequence from nucleotide 221 to 3004 of SEQ ID NO:5.

10. A nucleic acid molecule encoding a methyltransferase capable of catalyzing at least one of the reactions glycine to sarcosine, sarcosine to dimethyl glycine and dimethyl glycine to betaine.

11. A nucleic acid molecule comprising a nucleotide sequence selected form the group consisting of:

(a) a DNA sequence from nucleotide 2027 to 2722 of SEQ ID NO:1,
a DNA sequence from nucleotide 2006 to 3004 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c).

12. A nucleic acid molecule according to claim 11 which comprises the DNA sequence from nucleotide 2027 to 2722 of SEQ ID NO:1.

13. A nucleic acid molecule according to claim 11 which comprises the DNA sequence from nucleotide 2006 to 3004 of SEQ ID NO:5.

14. A methyltransferase encoded by a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b),

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c).

15. A methyltransferase according to claim 14 encoded by a nucleic acid molecule having the DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1.

16. A methyltransferase according to claim 14 encoded by a nucleic acid molecule having the DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5.

17. A methyltransferase according to claim 14 encoded by a nucleic acid molecule having the DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1.

18. A methyltransferase according to claim 14 encoded by a nucleic acid molecule having the DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5.

19. A methyltransferase comprising an amino acid sequence selected from the group consisting of:
(a) an amino acid sequence as depicted in SEQ ID NO:2, an amino acid sequence as depicted in SEQ ID NO:3, an amino acid sequence as depicted in SEQ ID NO:6, an amino acid sequence as depicted in SEQ ID NO:7,
(b) a fragment of an amino acid sequence as defined in (a) and
(c) a derivative of an amino acid sequence as defined in (a) and (b).

20. A methyltransferase according to claim 19 having the amino acid sequence depicted in SEQ ID NO:2.

21. A methyltransferase according to claim 19 having the amino acid sequence depicted in SEQ ID NO:3.

22. A methyltransferase according to claim 19 having the amino acid sequence depicted in SEQ ID NO:6.

23. A methyltransferase according to claim 19 having the amino acid sequence depicted in SEQ ID NO:7.

24. A methyltransferase according to claim 19 having the amino acid sequence depicted in SEQ ID NO:2 and SEQ ID NO:3, wherein the N-terminus of SEQ ID NO:3 is covalently joined to the C-terminus of SEQ ID NO:2.

25. A methyltransferase capable of catalyzing at least one of the reactions glycine to sarcosine, sarcosine to dimethyl glycine and dimethyl glycine to betaine.

26. A S-adenosyl methionine synthase encoded by a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 2027 to 2722 of SEQ ID NO:1,
a DNA sequence from nucleotide 2006 to 3004 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b),

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c).

27. A S-adenosyl methionine synthase encoded by a nucleic acid molecule having the DNA sequence from nucleotide 2027 to 2722 of SEQ ID NO:1.

28. A S-adenosyl methionine synthase encoded by a nucleic acid molecule having the DNA sequence from nucleotide 2006 to 3004 of SEQ ID NO:5.

29. A S-adenosyl methionine synthase comprising an amino acid sequence selected form the group consisting of:

(a) an amino acid sequence as depicted in SEQ ID NO:4 or
an amino acid sequence as depicted in SEQ ID NO:8,

(b) a fragment of an amino acid sequence as defined in (a) and

(c) a derivative of an amino acid sequence as defined in (a) and (b).

30. A S-adenosyl methionine synthase according to claim 26 having the amino acid sequence depicted in SEQ ID NO:4.

31. A S-adenosyl methionine synthase according to claim 26 having the amino acid sequence depicted in SEQ ID NO:8.

32. An expression vector comprising at least one nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c).

33. An expression vector according to claim 32 comprising the DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1.

34. An expression vector according to claim 32 comprising the DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5.

35. An expression vector according to claim 32 comprising the DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1.

36. An expression vector according to claim 32 comprising the DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5.

37. An expression vector according to claim 32 comprising the DNA sequence from nucleotide 208 to 1902 of SEQ ID NO:1.

38. An expression vector according to claim 32 comprising the DNA sequence from nucleotide 221 to 1867 of SEQ ID NO:5.

39. An expression vector comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

40. A recombinant organism transformed with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,

a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

41. A recombinant organism according to claim 40 which is transformed with an expression vector comprising the DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1.

42. A recombinant organism according to claim 40 which is transformed with an expression vector comprising the DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5.

43. A recombinant organism according to claim 40 which is transformed with an expression vector comprising the DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1.

44. A recombinant organism according to claim 40 which is transformed with an expression vector comprising the DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5.

45. A recombinant organism according to claim 40 which is transformed with an expression vector comprising the DNA sequence from nucleotide 208 to 1902 of SEQ ID NO:1.

46. A recombinant organism according to claim 40 which is transformed with an expression vector comprising the DNA sequence from nucleotide 221 to 1867 of SEQ ID NO:5.

47. A recombinant organism transformed with an nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c).

48. A recombinant organism transformed with an nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,

a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c).

49. A recombinant organism according to claim 40, wherein the organism is a bacteria.

50. A recombinant organism according to claim 49, wherein the bacteria is selected from the group consisting of E. coli, Bacillus, Corynebacteria, Pseudomonas lactic acid bacteria and Streptomyces.

51. A recombinant organism according to claim 40, wherein the organism is a yeast.

52. A recombinant organism according to claim 51, wherein the yeast is selected from the group consisting of Saccharomyces, Zygosaccharomyces, Pichia, Kluyveromyces, Candida and Hansenula.

53. A recombinant organism according to claim 40, wherein the organism is a fungus.

54. A recombinant organism according to claim 53, wherein the fungus is selected from the group consisting of Aspergillus, Trichoderma and Penicillium.

55. A recombinant organism according to claim 40, wherein the organism is a plant selected from the

group comprising cereals, legumes, oilseeds, vegetables, fruits, ornamentals and perennial trees.

56. A recombinant organism according to claim 55, wherein the plant is selected from the group consisting of lettuces, Capsicums, grasses, clovers, alfalfa, beans, sweet potatoes, cassava, yams, taro, groundnut, brassica, sugar beet, grapes, potato, tomato, rice, tobacco, rapeseed, maize, sorghum, cotton, soybean, barley, wheat, rye, canola, sunflower, linseed, pea, cucumber, carrot, ornamentals, perennial trees and fruits.

57. A recombinant organism according to claim 47, wherein the organism is a bacteria.

58. A recombinant organism according to claim 57, wherein the bacteria is selected from the group consisting of E. coli, Bacillus, Corynebacteria, Pseudomonas lactic acid bacteria and Streptomyces.

59. A recombinant organism according to claim 47, wherein the organism is a yeast.

60. A recombinant organism according to claim 59, wherein the yeast is selected from the group consisting of Saccharomyces, Zygosaccharomyces, Pichia, Kluyveromyces, Candida and Hansenula.

61. A recombinant organism according to claim 47, wherein the organism is a fungus.

62. A recombinant organism according to claim 61, wherein the fungus is selected from the group consisting of Aspergillus, Trichoderma and Penicillium.

63. A recombinant organism according to claim 47, wherein the organism is a plant selected from the group comprising cereals, legumes, oilseeds, vegetables, fruits, ornamentals and perennial trees.

64. A recombinant organism according to claim 63, wherein the plant is selected from the group consisting of lettuces, Capsicums, grasses, clovers, alfalfa, beans, sweet potatoes, cassava, yams, taro, groundnut, brassica, sugar beet, grapes, potato, tomato, rice, tobacco, rapeseed, maize, sorghum, cotton, soybean, barley, wheat, rye, canola, sunflower, linseed, pea, cucumber, carrot, ornamentals, perennial trees and fruits.

65. A recombinant organism according to claim 48, wherein the organism is a bacteria.

66. A recombinant organism according to claim 66, wherein the bacteria is selected from the group consisting of E. coli, Bacillus, Corynebacteria, Pseudomonas lactic acid bacteria and Streptomyces.

67. A recombinant organism according to claim 48, wherein the organism is a yeast.

68. A recombinant organism according to claim 67, wherein the yeast is selected from the group consisting of Saccharomyces, Zygosaccharomyces, Pichia, Kluyveromyces, Candida and Hansenula.

69. A recombinant organism according to claim 48, wherein the organism is a fungus.

70. A recombinant organism according to claim 69, wherein the fungus is selected from the group

consisting of Aspergillus, Trichoderma and Penicillium.

71. A recombinant organism according to claim 48, wherein the organism is a plant selected from the group comprising cereals, legumes, oilseeds, vegetables, fruits, ornamentals and perennial trees.

72. A recombinant organism according to claim 71, wherein the plant is selected from the group consisting of lettuces, Capsicums, grasses, clovers, alfalfa, beans, sweet potatoes, cassava, yams, taro, groundnut, brassica, sugar beet, grapes, potato, tomato, rice, tobacco, rapeseed, maize, sorghum, cotton, soybean, barley, wheat, rye, canola, sunflower, linseed, pea, cucumber, carrot, ornamentals, perennial trees and fruits.

73. A method for the production of a recombinant organism comprising the steps of transforming a host organism with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

74. A method for the production of a recombinant organism comprising the steps of transforming a host organism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

75. A method for the production of a recombinant organism comprising the steps of transforming a host organism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,

a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c)

76. A methyltransferase obtainable by culturing an organism transformed with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c)
and isolating the methyltransferase from said organism or the medium used to culture said organism.

77. A method for the production of a methyltransferase comprising the steps of culturing an organism transformed with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c),

and isolating the methyltransferase from said organism or the medium used to culture said organism.

78. A method for the production of sarcosine comprising the steps of culturing an organism transformed with a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c),
and isolating dimethyl glycine from said organism or the medium used to culture or process said organism.

79. A method for the production of sarcosine comprising the steps of culturing an organism transformed with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c),
and isolating sarcosine from said organism or the medium used to culture or process said organism.

80. A method for the production of sarcosine comprising the steps of culturing an organism transformed with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least

one a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c),

and isolating sarcosine from said organism or the medium used to culture or process said organism.

81. A method for the production of dimethyl glycine comprising the steps of culturing an organism transformed with a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c),

and isolating dimethyl glycine from said organism or the medium used to culture or process said organism.

82. A method for the production of dimethyl glycine comprising the steps of culturing an organism transformed with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c),
and isolating dimethyl glycine from said organism or the medium used to culture or process said organism.

83. A method for the production of dimethyl glycine comprising the steps of culturing an organism transformed with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c),
and isolating dimethyl glycine from said organism or the medium used to culture or process said organism.

84. A method for the production of betaine, comprising the steps of culturing an organism transformed with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c), and isolating betaine from said organism or the medium used to culture or process said organism.

85. A method for the production of betaine, comprising the steps of culturing an organism transformed with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c), and isolating betaine from said organism or the medium used to culture or process said organism.

86. A method for the production of betaine, comprising the steps of culturing an organism transformed with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c),
and isolating betaine from said organism or the medium used to culture or process said organism.

87. A method for increasing the intracellular concentration of sarcosine, dimethyl glycine or betaine in an organism comprising the steps of transforming an organism with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide

sequences defined in (a), (b)and (c), such that the nucleic acid molecule(s) is expressed.

88. A method for increasing the intracellular concentration of sarcosine, dimethyl glycine or betaine in an organism comprising the steps of transforming an organism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

89. A method for increasing the intracellular concentration of sarcosine, dimethyl glycine or betaine in an organism comprising the steps of transforming an organism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,

a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

90. A method for enhancing the salt tolerance of an organism comprising the steps of transforming an organism with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c), such that the nucleic acid molecule(s) is expressed.

91. A method for enhancing the salt tolerance of an organism comprising the steps of transforming an

organism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

92. A method for enhancing the salt tolerance of an organism comprising the steps of transforming an organism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

93. A method for enhancing the freezing or cold tolerance of an organism comprising the steps of transforming an organism with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c), such that the nucleic acid molecule(s) is expressed.

94. A method for enhancing the freezing or cold tolerance of an organism comprising the steps transforming an organism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

95. A method for enhancing the freezing or cold tolerance of an organism comprising the steps transforming an organism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

96. A method for enhancing the resistance of an organism to drought or water stress comprising the steps of transforming an organism with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c), such that the nucleic acid molecule(s) is expressed.

97. A method for enhancing the resistance of an organism to drought or water stress comprising the steps of transforming an organism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,

a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

98. A method for enhancing the resistance of an organism to drought or water stress comprising the steps of transforming an organism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

99. A method for enhancing the productivity or yield of an organism comprising the steps of transforming an organism with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c), such that the nucleic acid molecule(s) is expressed.

100. A method for enhancing the productivity or yield of an organism comprising the steps of transforming an organism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c).

101. A method for enhancing the productivity or yield of an organism comprising the steps of transforming an organism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c).

102. A method for inducing pathogenesis-related proteins in a plant comprising the steps of transforming a plant with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c), such that the nucleic acid molecule(s) is expressed.

103. A method for inducing pathogenesis-related proteins in a plant comprising the steps of transforming a plant with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

104. A method for inducing pathogenesis-related proteins in a plant comprising the steps of transforming a plant with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

105. A method for increasing the resistance of a plant to attack by pathogens comprising the steps of transforming a plant with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,

a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c), such that the nucleic acid molecule(s) is expressed.

106. A method for increasing the resistance of a plant to attack by pathogens comprising the steps of transforming a plant with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

107. A method for increasing the resistance of a plant to attack by pathogens comprising the steps of transforming a plant with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c).

108. A method for improving the nutritional value of a plant comprising the steps of transforming a plant with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c), such that the nucleic acid molecule(s) is expressed.

109. A method for improving the nutritional value of a plant comprising the steps of transforming a plant with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

110. A method for improving the nutritional value of a plant comprising the steps of transforming a plant with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and

at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

111. A method for enhancing the pH tolerance of a cultured microorganism comprising the steps of transforming a microorganism with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide

sequences defined in (a), (b) and (c), such that the nucleic acid molecule(s) is expressed.

112. A method for enhancing the pH tolerance of a cultured microorganism comprising the steps of transforming a microorganism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c).

113. A method for enhancing the pH tolerance of a cultured microorganism comprising the steps of transforming a microorganism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,

a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

114. A method for improving the viability of a cultured microorganism comprising the steps of transforming an microorganism with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c), such that the nucleic acid molecule(s) is expressed.

115. A method for improving the viability of a cultured microorganism comprising the steps of transforming

a microorganism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

116. A method for improving the viability of a cultured microorganism comprising the steps of transforming a microorganism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

117. A method for decreasing inclusion body formation in a microorganism expressing a heterologous protein comprising the steps of transforming a microorganism with a nucleic acid molecule capable of expressing a heterologous protein and transforming said microorganism with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c), such that the nucleic acid molecules are expressed.

118. A method for decreasing inclusion body formation in a microorganism expressing a heterologous protein comprising the steps of transforming a microorganism with a nucleic acid molecule capable of expressing a heterologous protein and

transforming said microorganism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

119. A method for decreasing inclusion body formation in a microorganism expressing a heterologous protein comprising the steps of transforming a microorganism with a nucleic acid molecule capable of expressing a heterologous protein and transforming said microorganism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,

a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

120. A method for increasing the stability of a heterologous protein expressed in a microorganism comprising the steps of transforming a microorganism with a nucleic acid molecule capable of expressing a heterologous protein and transforming said microorganism with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c), such that the nucleic acid molecules are expressed.

121. A method for increasing the stability of a heterologous protein expressed in a microorganism comprising the steps of transforming a microorganism with a nucleic acid molecule capable of expressing a heterologous protein and transforming said microorganism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

122. A method for increasing the stability of a heterologous protein expressed in a microorganism comprising the steps of transforming a microorganism with a nucleic acid molecule capable of expressing a heterologous protein and transforming said microorganism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

123. A method for increasing the production of a heterologous protein expressed in a microorganism comprising the steps of transforming a microorganism with a nucleic acid molecule capable of expressing a heterologous protein and transforming said microorganism with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide

sequences defined in (a), (b)and (c), such that the nucleic acid molecules are expressed.

124. A method for increasing the production of a heterologous protein expressed in a microorganism comprising the steps of transforming a microorganism with a nucleic acid molecule capable of expressing a heterologous protein and transforming said microorganism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

125. A method for increasing the production of a heterologous protein expressed in a microorganism comprising the steps of transforming a microorganism with a nucleic acid molecule capable of expressing a heterologous protein and transforming said microorganism with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of

intracellular glycine and at least one nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

126. An animal feed comprising a recombinant organism transformed with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

127. An animal feed comprising a recombinant organism transformed with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least one nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b) and (c).

128. An animal feed comprising a recombinant organism transformed with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular glycine and at least one nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

129. An animal feed ingredient comprising a recombinant organism transformed with at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

130. An animal feed ingredient comprising a recombinant organism transformed with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,

a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

131. An animal feed ingredient comprising a recombinant organism transformed with a nucleic acid molecule comprising a nucleotide sequence coding for an enzyme capable of increasing the amount of intracellular S-adenosyl methionine and at least nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,

(b) a nucleotide sequence which hybridizes to any of the sequences as defined in (a),

(c) a fragment of a nucleotide sequence as defined in (a) and (b), and

(d) a nucleotide sequence which is degenerate as a result of the genetic code to any of the nucleotide sequences defined in (a), (b)and (c).

132. A DNA probe for use in identifying and cloning a nucleic acid molecule encoding a methyltransferase

comprising at least 15 nucleotides of a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5.

133. A method for identifying and cloning a nucleic acid molecule encoding a methyltransferase comprising the steps of hybridizing a probe consisting of at least 15 nucleotides of a nucleotide sequence selected from the group consisting of:

(a) a DNA sequence from nucleotide 208 to 1047 of SEQ ID NO:1,
a DNA sequence from nucleotide 221 to 1024 of SEQ ID NO:5,
a DNA sequence from nucleotide 1048 to 1902 of SEQ ID NO:1,
a DNA sequence from nucleotide 1031 to 1867 of SEQ ID NO:5,
with a sample containing nucleic acid of an organism, detecting a nucleic acid molecule in said sample which hybridizes to said probe and isolating said detected nucleic acid molecule.

134. A method for the purification of a methyl-transferase capable of catalyzing the conversion of glycine to dimethyl glycine comprising the steps of subjecting a sample comprising the methyl-transferase to a matrix containing adenosine, binding said methyltransferase to said matrix and eluting said methyltransferase from said matrix.

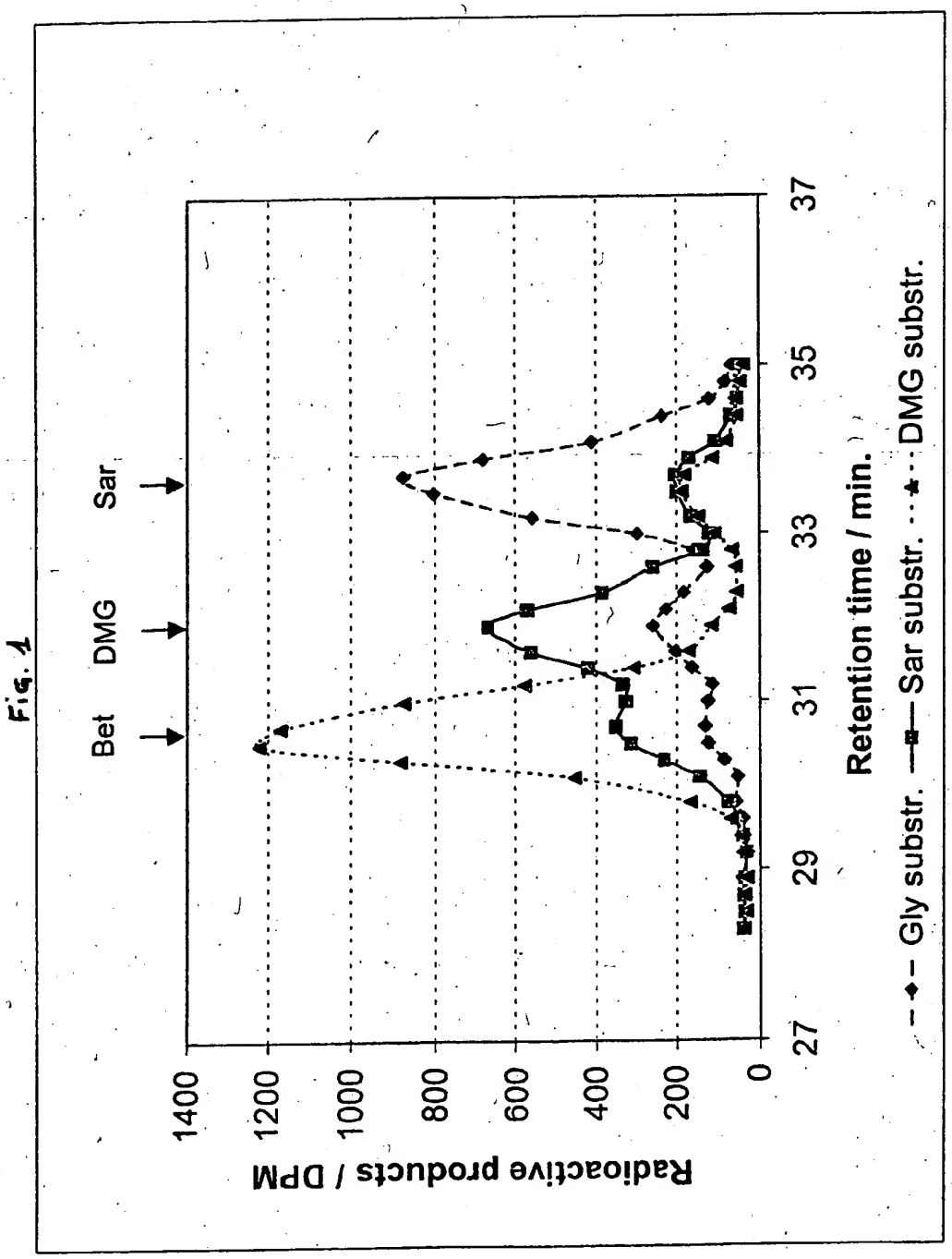


Fig. 1
Bet
DMG
Sar
Radioactive products / DPM
1400
1200
1000
800
600
400
200
0
27
29
31
33
35
37
Retention time / min.
Gly substr.
Sar substr.
DMG substr.

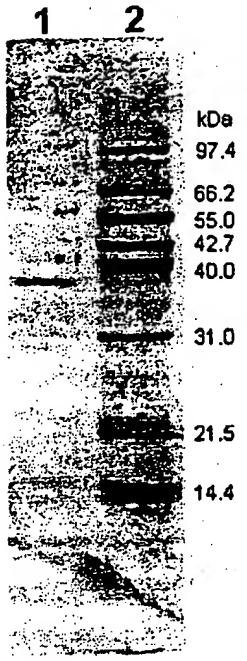


Fig. 2A

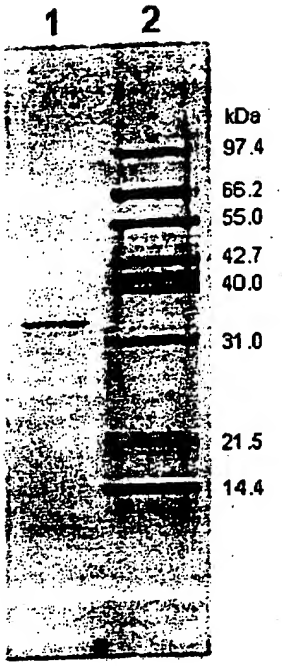


Fig. 2B

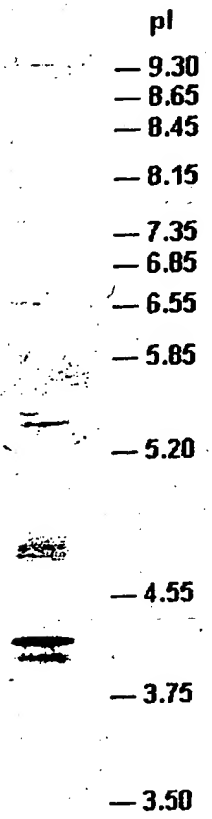


Fig. 3

Fig. 4

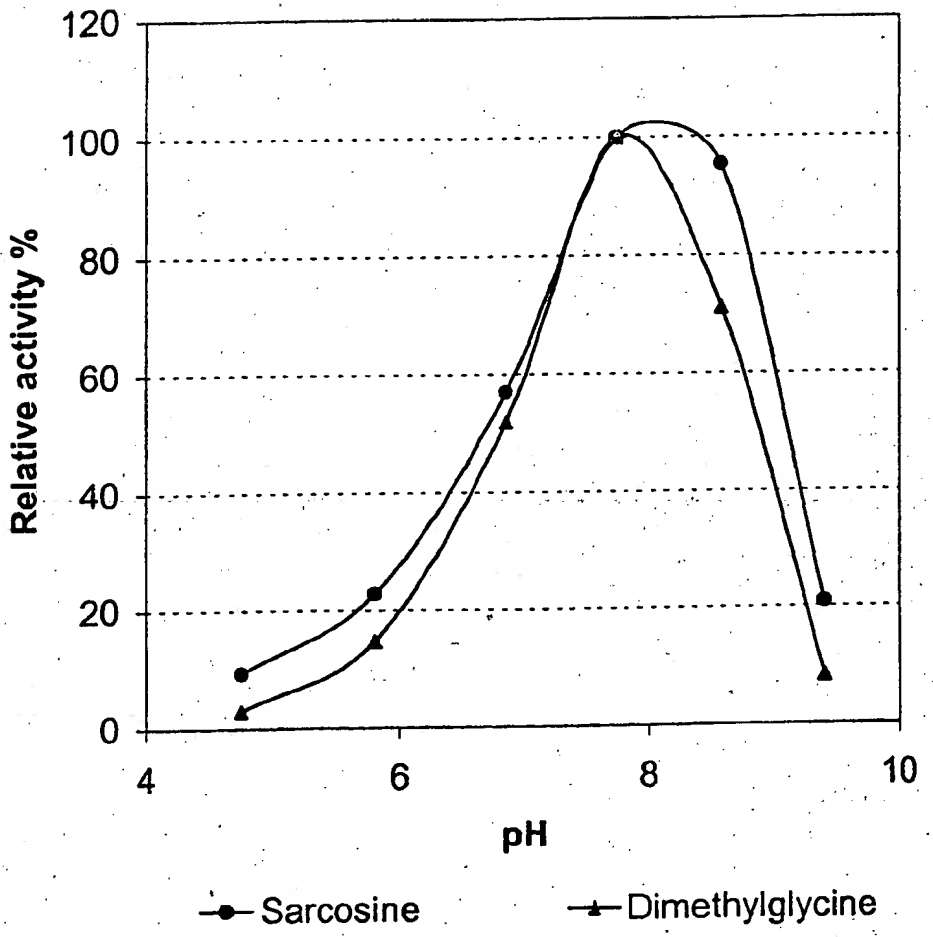

120
100
80
60
40
20
0
Relative activity %
4
6
8
10
pH
Sarcosine
Dimethylglycine

Fig. 5

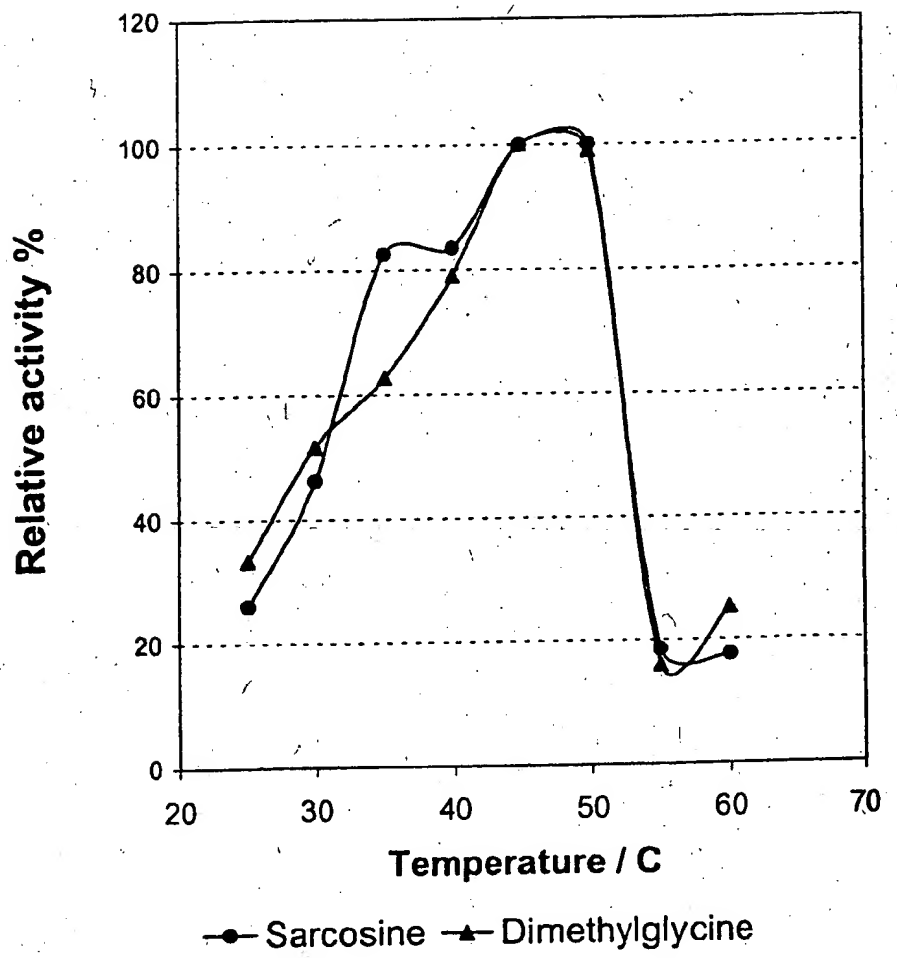

120
100
80
60
40
20
0
Relative activity %
20
30
40
50
60
70
Temperature / C
Sarcosine
Dimethylglycine

Fig. 6

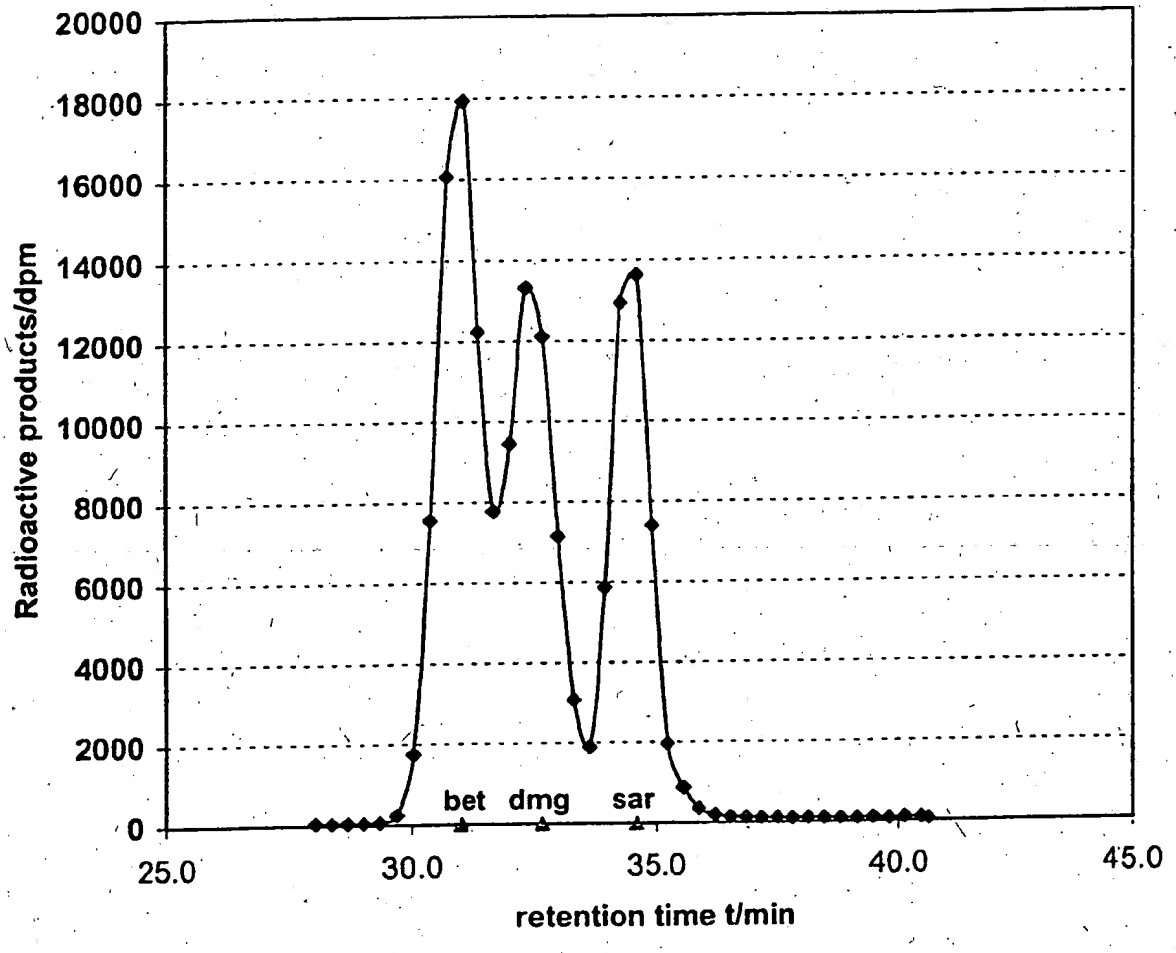

20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
Radioactive products/dpm
bet
dmg
sar
25.0
30.0
35.0
40.0
45.0
retention time t/min

Fig. 7

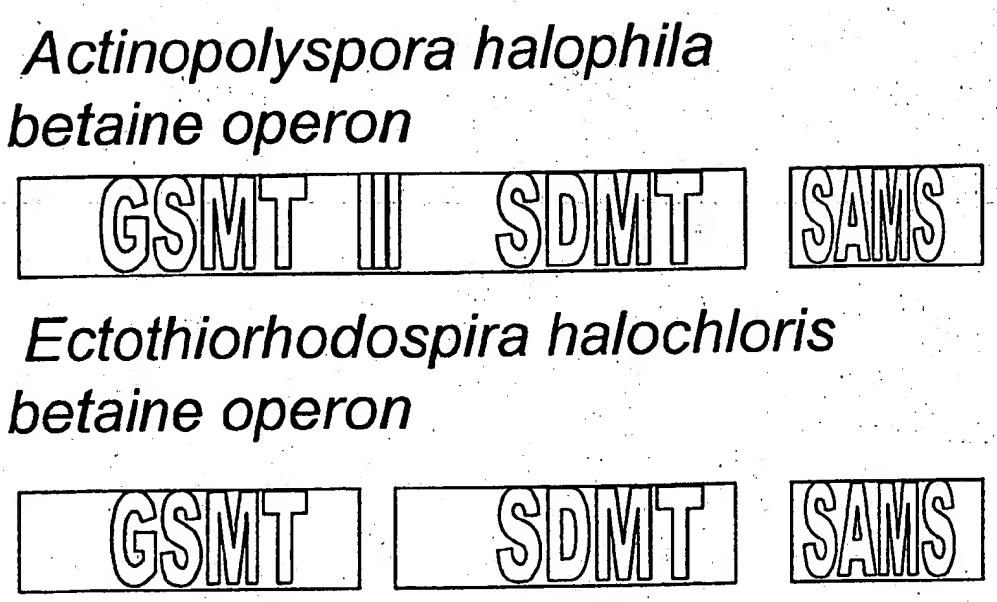

Actinopolyspora halophila
betaine operon
GSMT III SDMT SAMS
Ectothiorhodospira halochloris
betaine operon
GSMT SDMT SAMS

Figure 8

```
GCAGGCTCGG TGCCACCAAT ACAGGGCAGC CCCCTGGCCA GAAACACGCC CCGTAGTTAC      60

GGAGGCGCAT GCCATAGATA ATGAACTTGC TACAAGAGGC ACTTGACATT CAGTTGTCCA     120

AGCGGCGCCG GCGATCAAGC ACACGATCTG AATTCACAGC GTATGGGTAT ACCGTAGCGC     180

                                                    primer 1.
GCGTTCGCTA CACCAACCAT TCGCCAAGCT GAGGTGATAT ATG AAT ACG ACT ACT        235
                                               Met Asn Thr Thr Thr
                                                1               5
                                                 ↳ GSMT
GAG CAG GAC TTC GGA GCG GAC CCA ACA AAA GTC CGA GAT ACC GAC CAC        283
Glu Gln Asp Phe Gly Ala Asp Pro Thr Lys Val Arg Asp Thr Asp His
               10                  15                  20

TAC ACG GAA GAG TAT GTT GAC GGG TTT GTC GAC AAA TGG GAC GAC TTG        331
Tyr Thr Glu Glu Tyr Val Asp Gly Phe Val Asp Lys Trp Asp Asp Leu
           25                  30                  35

ATT GAT TGG GAT AGC CGC GCG AAG AGT GAA GGT GAT TTC TTT ATT CAA        379
Ile Asp Trp Asp Ser Arg Ala Lys Ser Glu Gly Asp Phe Phe Ile Gln
       40                  45                  50

GAG TTG AAG AAG CGT GGT GCC ACA CGC ATC CTC GAT GCC GCA ACA GGT        427
Glu Leu Lys Lys Arg Gly Ala Thr Arg Ile Leu Asp Ala Ala Thr Gly
   55                  60                  65

ACC GGC TTC CAC TCA GTG CGT TTG CTC GAG GCG GGT TTT GAT GTC GTC        475
Thr Gly Phe His Ser Val Arg Leu Leu Glu Ala Gly Phe Asp Val Val
70                  75                  80                  85

AGT GCT GAT GGC AGT GCC GAG ATG CTC GCC AAA GCC TTC GAG AAT GGG        523
Ser Ala Asp Gly Ser Ala Glu Met Leu Ala Lys Ala Phe Glu Asn Gly
               90                  95                 100

CGT AAG CGT GGC CAT ATC CTA CGC ACG GTT CAG GTC GAC TGG CGG TGG        571
Arg Lys Arg Gly His Ile Leu Arg Thr Val Gln Val Asp Trp Arg Trp
           105                 110                 115

CTG AAT CGC GAT ATC CAC GGT CGT TAT GAC GCC ATT ATC TGC CTT GGC        619
Leu Asn Arg Asp Ile His Gly Arg Tyr Asp Ala Ile Ile Cys Leu Gly
       120                 125                 130

AAC TCG TTT ACC CAC CTG TTT AAT GAA AAA GAT CGG CGC AAG ACC CTG        667
Asn Ser Phe Thr His Leu Phe Asn Glu Lys Asp Arg Arg Lys Thr Leu
   135                 140                 145

GCA GAG TTC TAT TCC GCG CTG AAC CCG GAA GGG GTG CTG ATA TTG GAT        715
Ala Glu Phe Tyr Ser Ala Leu Asn Pro Glu Gly Val Leu Ile Leu Asp
150                 155                 160                 165

CAG CGC AAT TAC GAC GGC ATC CTC GAT CAC GGC TAT GAC TCA AGC CAC        763
Gln Arg Asn Tyr Asp Gly Ile Leu Asp His Gly Tyr Asp Ser Ser His
               170                 175                 180

TCC TAC TAC TAT TGC GGT GAG GGC GTC TCT GTC TAT CCT GAG CAC GTT        811
Ser Tyr Tyr Tyr Cys Gly Glu Gly Val Ser Val Tyr Pro Glu His Val
           185                 190                 195

GAT GAT GGC CTA GCC CGC TTC AAG TAT GAA TTC AAC GAC GGT TCA ACT        859
Asp Asp Gly Leu Ala Arg Phe Lys Tyr Glu Phe Asn Asp Gly Ser Thr
       200                 205                 210
```

```
TAC TTC CTG AAC ATG TTC CCG CTG CGC AAG GAT TAC ACG CGC CGC CTA      907
Tyr Phe Leu Asn Met Phe Pro Leu Arg Lys Asp Tyr Thr Arg Arg Leu
    215                 220                 225

ATG CAT GAG GTG GGT TTC CAA AAA ATC GAT ACC TAC GGC GAC TTC AAG      955
Met His Glu Val Gly Phe Gln Lys Ile Asp Thr Tyr Gly Asp Phe Lys
230                 235                 240                 245

GCA ACC TAC CGC GAT GCA GAT CCG GAT TTC TTT ATT CAT GTC GCC GAG     1003
Ala Thr Tyr Arg Asp Ala Asp Pro Asp Phe Phe Ile His Val Ala Glu
                250                 255                 260
          primer 2                          primer 3
AAG GAA TAT CGG GAG GAG GAC TGATAT ATG GCG ACG CGC TAC GAC GAT      1051
Lys Glu Tyr Arg Glu Glu Asp  *     Met Ala Thr Arg Tyr Asp Asp
            265                      1                   5
                    GSMT ←            → SDMT
CAA GCC ATA GAA ACG GCG AGG CAG TAC TAT AAC AGT GAG GAT GCC GAT     1099
Gln Ala Ile Glu Thr Ala Arg Gln Tyr Tyr Asn Ser Glu Asp Ala Asp
        10                  15                  20

AAC TTC TAC GCC ATA ATA TGG GGT GGT GAG GAT ATC CAT ATC GGT TTG     1147
Asn Phe Tyr Ala Ile Ile Trp Gly Gly Glu Asp Ile His Ile Gly Leu
    25                  30                  35

TAC AAC GAT GAT GAA GAG CCG ATA GCA GAT GCC AGC AGG CGT ACT GTA     1195
Tyr Asn Asp Asp Glu Glu Pro Ile Ala Asp Ala Ser Arg Arg Thr Val
40                  45                  50                  55

GAG CGC ATG TCG TCA CTG TCC CGG CAG CTT GGC CCA GAT AGC TAT GTC     1243
Glu Arg Met Ser Ser Leu Ser Arg Gln Leu Gly Pro Asp Ser Tyr Val
                60                  65                  70

CTC GAC ATG GGG GCT GGT TAC GGA GGG TCT GCC CGT TAC CTT GCG CAT     1291
Leu Asp Met Gly Ala Gly Tyr Gly Gly Ser Ala Arg Tyr Leu Ala His
            75                  80                  85

AAG TAT GGC TGT AAG GTG GCG GCG CTC AAT CTC TCC GAA CGT GAG AAT     1339
Lys Tyr Gly Cys Lys Val Ala Ala Leu Asn Leu Ser Glu Arg Glu Asn
        90                  95                  100

GAG CGT GAT CGG CAG ATG AAT AAG GAG CAA GGA GTC GAT CAC CTG ATT     1387
Glu Arg Asp Arg Gln Met Asn Lys Glu Gln Gly Val Asp His Leu Ile
    105                 110                 115

GAG GTT GTC GAT GCT GCC TTT GAA GAT GTC CCC TAT GAT GAT GGG GTG     1435
Glu Val Val Asp Ala Ala Phe Glu Asp Val Pro Tyr Asp Asp Gly Val
120                 125                 130                 135

TTT GAC CTG GTC TGG TCG CAG GAC TCC TTC CTC CAT AGT CCG GAT AGG     1483
Phe Asp Leu Val Trp Ser Gln Asp Ser Phe Leu His Ser Pro Asp Arg
                140                 145                 150

GAA CGC GTA CTG CGG GAG GCT AGT AGA GTT CTG CGT TCT GGG GGC GAG     1531
Glu Arg Val Leu Arg Glu Ala Ser Arg Val Leu Arg Ser Gly Gly Glu
            155                 160                 165

TTT ATC TTC ACC GAT CCA ATG CAG GCG GAT GAT TGC CCC GAG GGG GTT     1579
Phe Ile Phe Thr Asp Pro Met Gln Ala Asp Asp Cys Pro Glu Gly Val
        170                 175                 180

ATT CAA CCG ATC CTC GAT AGG ATC CAC CTC GAG ACG ATG GGC ACG CCT     1627
Ile Gln Pro Ile Leu Asp Arg Ile His Leu Glu Thr Met Gly Thr Pro
    185                 190                 195
```

```
AAT TTC TAT CGC CAG ACC CTT AGG GAT CTG GGC TTT GAG GAA ATC ACC      1675
Asn Phe Tyr Arg Gln Thr Leu Arg Asp Leu Gly Phe Glu Glu Ile Thr
200                 205                 210                 215

TTC GAA GAC CAC ACC CAT CAG TTG CCG CGC CAT TAC GGG CGC GTG CGC      1723
Phe Glu Asp His Thr His Gln Leu Pro Arg His Tyr Gly Arg Val Arg
                220                 225                 230

CGT GAG CTA GAT CGT CGA GAA GGT GAA CTG CAG GGG CAT GTC TCG GCA      1771
Arg Glu Leu Asp Arg Arg Glu Gly Glu Leu Gln Gly His Val Ser Ala
            235                 240                 245

GAG TAC ATC GAA CGC ATG AAA AAT GGC CTT GAC CAC TGG GTC AAT GGC      1819
Glu Tyr Ile Glu Arg Met Lys Asn Gly Leu Asp His Trp Val Asn Gly
        250                 255                 260
                                        primer 4
GGT AAC AAG GGT TAT CTC ACC TGG GGT ATC TTT TAC TTC CGC AAA GGG      1867
Gly Asn Lys Gly Tyr Leu Thr Trp Gly Ile Phe Tyr Phe Arg Lys Gly
    265                 270                 275
                                                    SDMT

TGATCACTAG AGCGCTAAAT CGCAGGCGCG GTAATTGTGC CGCGCTTCGG GCGCATATGT    1927
*
TTTAGCCCAG TGGATTTGCA CATAGAAAAG AGTGAATTCT GGTTTAGCGC TCGAATAAAC    1987

TATAGCAAAG GAGAAACT ATG ACT AAG CGA TAT CTA TTT ACC TCT GAG TCG      2038
                    Met Thr Lys Arg Tyr Leu Phe Thr Ser Glu Ser
                     1               5                  10
                      SAMS
GTC TCT GAA GGC CAC CCG GAC AAA ATG GCC GAC CAG ATT TCG GAT GCA      2086
Val Ser Glu Gly His Pro Asp Lys Met Ala Asp Gln Ile Ser Asp Ala
            15                  20                  25

CTG CTC GAT GAG TTC CTG CGC CAG GAT CCA AAG TCG CGG GTG GCT GCA      2134
Leu Leu Asp Glu Phe Leu Arg Gln Asp Pro Lys Ser Arg Val Ala Ala
        30                  35                  40

GAG ACG ATG ATT CAG ACC GGC ATG GTT GTA GTT GCC GGC GAA ATT AAG      2182
Glu Thr Met Ile Gln Thr Gly Met Val Val Val Ala Gly Glu Ile Lys
    45                  50                  55

AGC AAT GCC AAG ATC AAT GTT GAG CCG CTG GTA CGT GAA GTT GTC CGC      2230
Ser Asn Ala Lys Ile Asn Val Glu Pro Leu Val Arg Glu Val Val Arg
60                  65                  70                  75

GAT ATC GGC TAC ACC AGC TCA GAT ATG GGC TTT GAT GCC GAC ACC TGT      2278
Asp Ile Gly Tyr Thr Ser Ser Asp Met Gly Phe Asp Ala Asp Thr Cys
                80                  85                  90

GCC GTA CTC AAC GCC CTC GGC GAG CAG TCC CCC GAC ATC AAT CAA GGC      2326
Ala Val Leu Asn Ala Leu Gly Glu Gln Ser Pro Asp Ile Asn Gln Gly
            95                  100                 105

GTT GAC CGG GAA GAG GAA GAG GAG CAG GGC GCC GGT GAC CAA GGA CTG      2374
Val Asp Arg Glu Glu Glu Glu Glu Gln Gly Ala Gly Asp Gln Gly Leu
        110                 115                 120

ATG TTC GGT TAC GCC ACC AAT GAG ACC GAC GTC CTC ATG CCG GCA GCG      2422
Met Phe Gly Tyr Ala Thr Asn Glu Thr Asp Val Leu Met Pro Ala Ala
    125                 130                 135

ATC CAC TAC TCG CAC CTG CTG GTT AAG CGC CAA TCC GAA GTC CGT AAC      2470
Ile His Tyr Ser His Leu Leu Val Lys Arg Gln Ser Glu Val Arg Asn
140                 145                 150                 155
```

```
TCC AAG AAG CTG CCG TGG TTG CGC CCG GAC GCT AAG AGC CAG GTC ACC      2518
Ser Lys Lys Leu Pro Trp Leu Arg Pro Asp Ala Lys Ser Gln Val Thr
                160                 165                 170

TTC AAG TAC GAA GGC GAT AAG ATC GTC GGT TGC GAT GCG GTG GTG CTC      2566
Phe Lys Tyr Glu Gly Asp Lys Ile Val Gly Cys Asp Ala Val Val Leu
            175                 180                 185

TCC ACG CAG CAT GAC GAG ACC GTT GAT CAG AAG ACC GTC CAC GAG GGC      2614
Ser Thr Gln His Asp Glu Thr Val Asp Gln Lys Thr Val His Glu Gly
        190                 195                 200

GTG ATG GAA GAG ATC ATC AAG CCC ATC CTC GGT GAC ACC GGC TGG CTG      2662
Val Met Glu Glu Ile Ile Lys Pro Ile Leu Gly Asp Thr Gly Trp Leu
    205                 210                 215

ACC AAC GAG ACC AAG TAC CAC ATT AAC CCG ACC GGC CGC TTC GTT ACT      2710
Thr Asn Glu Thr Lys Tyr His Ile Asn Pro Thr Gly Arg Phe Val Thr
220                 225                 230                 235

GGT GGT CCG TTG GGT GAC TGT GGC TTG ACC GGC CGT AAG ATC ATC GTC      2758
Gly Gly Pro Leu Gly Asp Cys Gly Leu Thr Gly Arg Lys Ile Ile Val
                240                 245                 250

GAT ACC TAC GGC GGT ATG GGC CGT CAC GGC GGC GGT GCC TTC TCC GGC      2806
Asp Thr Tyr Gly Gly Met Gly Arg His Gly Gly Gly Ala Phe Ser Gly
            255                 260                 265

AAA GAT CCG TCT AAG GTC GAT CGC TCG GCG GCT TAT GTT GGC CGT TAT      2854
Lys Asp Pro Ser Lys Val Asp Arg Ser Ala Ala Tyr Val Gly Arg Tyr
        270                 275                 280

GTA GCT AAG AAC ATA GTG GCT GCC GGT CTG GCC GAC CGC TGT GAG GTG      2902
Val Ala Lys Asn Ile Val Ala Ala Gly Leu Ala Asp Arg Cys Glu Val
    285                 290                 295

CAG CTC TCC TAC GCC ATC GGT GTT GCC GAG CCG ACC TCG GTC AAT GTC      2950
Gln Leu Ser Tyr Ala Ile Gly Val Ala Glu Pro Thr Ser Val Asn Val
300                 305                 310                 315

GAA ACC TTT GGT ACC GGC AAG GTC GAA GAA GAG CTC GCG CAC GAT GGC      2998
Glu Thr Phe Gly Thr Gly Lys Val Glu Glu Glu Leu Ala His Asp Gly
                320                 325                 330

CAG CCC TAAGCCGTGG CCGCTGCGCG AGCCCTGGTA GCCGGCGCGA TCGAACGCCG      3054
Gln Pro *

SAMS
CAAAGATATG TGCCTGCTGC TCTGGGGGGA TCCCCGGACC GGTATCGCGC ACTTGGAAGG   3114

TGATTCGGTC ATCGAGCTCC ACCGCGGTGG CGGCCGCTCT AGAACTAGTG GATCCCCCGG   3174

GCTGCAGGAA TTCGATATCA AGCTTATCGA TACCGTCGAC CTCGA                    3219
```

Figure 9

```
AATCTGCGCC GGGATCGGAC GGCAAAGGTG CCGACCGAGT GTGCTGATCG ATGCCAGCGC      60

GCGGCCCGCT AGTTCGGCTG CCCGCATCGG TCCGGCGTGC CCCTTGTGTG AGGTGACCCG     120

GCGCGCCCGT GAAAGCTCGG GGCGGCGTAA TTCGCCGTGG TTCCGAGCGA TGCAGTGGCG     180
                                                  primer 7
                                     ____________________________________
CTTGACCAAC GCGAGTGGGA GCTGAGT ATG ACC AAG AGC GTG GAC GAT CTT         231
                              Met Thr Lys Ser Val Asp Asp Leu
                              1               5
                              └──► GSMT
GCC CGT GGT GAC CAG GCC GGG GAC GAG CAG GAC CCG GTG CAC CGC GAG       279
Ala Arg Gly Asp Gln Ala Gly Asp Glu Gln Asp Pro Val His Arg Glu
    10                  15                  20

CAG CAG ACG TTC GGC GAC AAT CCG TTG GAA GTA CGC GAC ACT GAT CAC       327
Gln Gln Thr Phe Gly Asp Asn Pro Leu Glu Val Arg Asp Thr Asp His
25                  30                  35                  40

TAC ATG CAT GAG TAC GTC GGT GGT TTT GTC GAC AAG TGG GAC GAT CTG       375
Tyr Met His Glu Tyr Val Gly Gly Phe Val Asp Lys Trp Asp Asp Leu
                45                  50                  55

ATC GAT TGG AAG AAG CGC TAC GAA AGC GAG GGC AGC TTC TTC ATC GAC       423
Ile Asp Trp Lys Lys Arg Tyr Glu Ser Glu Gly Ser Phe Phe Ile Asp
            60                  65                  70

CAA TTG CGC GCA CGC GGT GTC GAG ACC GTG CTG GAC GCG GCG GCC GGG       471
Gln Leu Arg Ala Arg Gly Val Glu Thr Val Leu Asp Ala Ala Ala Gly
        75                  80                  85

ACC GGT TTC CAC TCG GTC CGG TTG CTC GAG GAG GGG TTT GAG ACC GTC       519
Thr Gly Phe His Ser Val Arg Leu Leu Glu Glu Gly Phe Glu Thr Val
    90                  95                  100

AGC GCG GAC GGC AGC CCG CAG ATG CTG GCC AAG GCC TTC AGT AAC GGA       567
Ser Ala Asp Gly Ser Pro Gln Met Leu Ala Lys Ala Phe Ser Asn Gly
105                 110                 115                 120

CTG GCC TAC AAC GGT CAC ATT CTG CGT GTG GTC AAC GCG GAC TGG CGT       615
Leu Ala Tyr Asn Gly His Ile Leu Arg Val Val Asn Ala Asp Trp Arg
                125                 130                 135

TGG CTC AAC CGT GAC GTG CAC GGT GAA TAC GAC GCG ATC ATT TGC CTG       663
Trp Leu Asn Arg Asp Val His Gly Glu Tyr Asp Ala Ile Ile Cys Leu
            140                 145                 150

GGC AAC TCC TTT ACC CAC CTG TTC TCG GAG CGG GAC CGC CGC AAG ACG       711
Gly Asn Ser Phe Thr His Leu Phe Ser Glu Arg Asp Arg Arg Lys Thr
        155                 160                 165

CTG GCT GAG TTC TAC GCG ATG CTC AAG CAC GAC GGT GTC CTG ATC ATC       759
Leu Ala Glu Phe Tyr Ala Met Leu Lys His Asp Gly Val Leu Ile Ile
    170                 175                 180

GAC CAG CGA AAC TAC GAC TCC ATT CTT GAC ACC GGC TTC TCC AGT AAG       807
Asp Gln Arg Asn Tyr Asp Ser Ile Leu Asp Thr Gly Phe Ser Ser Lys
185                 190                 195                 200

CAC ACG TAT TAC TAC GCC GGT GAG GAC GTT TCC GCG GAG CCC GAC CAC       855
His Thr Tyr Tyr Tyr Ala Gly Glu Asp Val Ser Ala Glu Pro Asp His
                205                 210                 215
```

```
ATC GAC GAC GGG CTG GCG CGG TTC AAG TAC ACG TTC CCG GAC AAG TCC      903
Ile Asp Asp Gly Leu Ala Arg Phe Lys Tyr Thr Phe Pro Asp Lys Ser
          220                 225                 230

GAA TTC TTC CTG AAC ATG TAC CCG CTG CGG AAA GAC TAC ATG CGG CGG      951
Glu Phe Phe Leu Asn Met Tyr Pro Leu Arg Lys Asp Tyr Met Arg Arg
      235                 240                 245

CTC ATG CGT GAG GTC GGT TTC CAA AGG ATT GAC ACC TAC GGT GAT TTC      999
Leu Met Arg Glu Val Gly Phe Gln Arg Ile Asp Thr Tyr Gly Asp Phe
  250                 255                 260

CAG GAA ACT TAC GGT GAA GAC GAG CCC GAC TTC TAC ATC CAC GTC GCG     1047
Gln Glu Thr Tyr Gly Glu Asp Glu Pro Asp Phe Tyr Ile His Val Ala
265       primer 5  270                 275                 280
                                                      GSMT
GAG AAG AGC TAC CGC ACC GAG GAC GAG TTC GTC GAC ATG TAC TCG AAC     1095
Glu Lys Ser Tyr Arg Thr Glu Asp Glu Phe Val Asp Met Tyr Ser Asn
 1                   5                  10                  15
 SDMT
GCG GTG CAC ACC GCG CGG GAC TAC TAC AAC TCC GAG GAC GCG GAC AAC     1143
Ala Val His Thr Ala Arg Asp Tyr Tyr Asn Ser Glu Asp Ala Asp Asn
          20                  25                  30

TTC TAC TAC CAC GTC TGG GGC GGC AAC GAC ATC CAC GTC GGG CTG TAC     1191
Phe Tyr Tyr His Val Trp Gly Gly Asn Asp Ile His Val Gly Leu Tyr
      35                  40                  45

CAG ACA CCG CAG GAG GAC ATC GCC ACC GCC AGT GAG CGC ACT GTC CAG     1239
Gln Thr Pro Gln Glu Asp Ile Ala Thr Ala Ser Glu Arg Thr Val Gln
  50                  55                  60

CGG ATG GCG GGC AAG GTC GAC ATC AGC CCC GAA ACC AGG ATT CTG GAT     1287
Arg Met Ala Gly Lys Val Asp Ile Ser Pro Glu Thr Arg Ile Leu Asp
65                  70                  75                  80

CTC GGT GCC GGC TAC GGC GGA GCC GCG CGG TAC CTG GCC AGG ACC TAC     1335
Leu Gly Ala Gly Tyr Gly Gly Ala Ala Arg Tyr Leu Ala Arg Thr Tyr
                85                  90                  95

GGC TGC CAC GTC ACC TGC CTC AAC CTC AGC GAG GTG GAG AAC CAG CGC     1383
Gly Cys His Val Thr Cys Leu Asn Leu Ser Glu Val Glu Asn Gln Arg
          100                 105                 110

AAC CGC GAG ATC ACT CGC GCC GAG GGG CTC GAG CAC CTG ATC GAG GTG     1431
Asn Arg Glu Ile Thr Arg Ala Glu Gly Leu Glu His Leu Ile Glu Val
      115                 120                 125

ACC GAC GGT TCC TTC GAG GAT CTC CCC TAC CAG GAC AAC GCG TTC GAC     1479
Thr Asp Gly Ser Phe Glu Asp Leu Pro Tyr Gln Asp Asn Ala Phe Asp
  130                 135                 140

GTG GTC TGG TCG CAG GAC TCC TTC CTC CAC AGC GGT GAC CGC AGC AGG     1527
Val Val Trp Ser Gln Asp Ser Phe Leu His Ser Gly Asp Arg Ser Arg
145                 150                 155                 160

GTC ATG GAA GAG GTG ACC CGG GTC CTC AAG CCG AAG GGT TCG GTG CTG     1575
Val Met Glu Glu Val Thr Arg Val Leu Lys Pro Lys Gly Ser Val Leu
                165                 170                 175

TTC ACC GAT CCG ATG GCG TCC GAC TCG GCG AAG AAG AAC GAG CTC GGC     1623
Phe Thr Asp Pro Met Ala Ser Asp Ser Ala Lys Lys Asn Glu Leu Gly
          180                 185                 190
```

```
CCC ATC CTG GAC AGG CTG CAC CTG GAC TCG CTC GGC TCG CCC GGT TTC      1671
Pro Ile Leu Asp Arg Leu His Leu Asp Ser Leu Gly Ser Pro Gly Phe
        195                 200                 205

TAC CGG AAG GAG CTG ACT CGT CTC GGG CTG CAG AAC ATC GAG TTC GAG      1719
Tyr Arg Lys Glu Leu Thr Arg Leu Gly Leu Gln Asn Ile Glu Phe Glu
    210                 215                 220

GAC CTC AGC GAA TAC CTG CCC GTC CAC TAC GGC CGG GTT CTG GAA GTG      1767
Asp Leu Ser Glu Tyr Leu Pro Val His Tyr Gly Arg Val Leu Glu Val
225                 230                 235                 240

CTG GAG AGC CGG GAG AAC GAG CTC GCC GGC TTC ATC GGC GAG GAG TAC      1815
Leu Glu Ser Arg Glu Asn Glu Leu Ala Gly Phe Ile Gly Glu Glu Tyr
                245                 250                 255

CGA GCT CAC ATG AAG ACC GGG CTG CGC AAC TGG GTG CAG GCC GGC AAT      1863
Arg Ala His Met Lys Thr Gly Leu Arg Asn Trp Val Gln Ala Gly Asn
            260                 265                 270
                                primer 6
GGC GGG AGC CTG GCC TGG GGC ATC ATC CAC GCC AGG GCA TGACAGCCGT       1912
Gly Gly Ser Leu Ala Trp Gly Ile Ile His Ala Arg Ala  *
        275                 280                 285
                                            SDMT
CGCGTACGAA GAACGATCAG AAGCAGTTAG CAGTGAGGTG AAGATCAGCC GTGACTGAGA    1972

TGAACCGCAG GTTGTTCACC AAGTGAGTCC GTGACCGAGG GCCACCCGGA CAAG ATG      2029
                                                          Met
                                                           1
                                                            SAMS
GCC GAC TCG ATC AGC GAC GCG ATC CTG GAC GCG ATG CTG GCT CAG GAC      2077
Ala Asp Ser Ile Ser Asp Ala Ile Leu Asp Ala Met Leu Ala Gln Asp
        5                   10                  15

CCC CGC TCC CGC GTG GCC ATG GAG ACC ATG ATC ACC ACC GGG CAG GTG      2125
Pro Arg Ser Arg Val Ala Met Glu Thr Met Ile Thr Thr Gly Gln Val
    20                  25                  30

CAC CTG GCC GGT GAG GTG ACC ACC GAG GCC GAC GTC GAC CTG CCC GCG      2173
His Leu Ala Gly Glu Val Thr Thr Glu Ala Asp Val Asp Leu Pro Ala
35                  40                  45

ATC GTG CGG GAG AAG GTC CTC GAG ATC GGC TAC GAC AAC TCG GCC AAG      2221
Ile Val Arg Glu Lys Val Leu Glu Ile Gly Tyr Asp Asn Ser Ala Lys
50                  55                  60                  65

GGC TTC GAC GGA GAC TCC TGC GGC ATC AAC GTC TCC ATC GAC GCG CAG      2269
Gly Phe Asp Gly Asp Ser Cys Gly Ile Asn Val Ser Ile Asp Ala Gln
                70                  75                  80

TCC CCG GAC ATC GGC CAG GGC GTG GAC TCC GCT CAC GAG TCC CGC GTC      2317
Ser Pro Asp Ile Gly Gln Gly Val Asp Ser Ala His Glu Ser Arg Val
            85                  90                  95

GAG GGT GCC ATC GAC GAG ATC GCC AGT CAG GGC GCC GGC GAC CAG GGC      2365
Glu Gly Ala Ile Asp Glu Ile Ala Ser Gln Gly Ala Gly Asp Gln Gly
        100                 105                 110

CTG ATG TTC GGT TAC GCC ACC AGC GAG ACC GAC GAG CTC ATG CCG CTG      2413
Leu Met Phe Gly Tyr Ala Thr Ser Glu Thr Asp Glu Leu Met Pro Leu
    115                 120                 125

CCG ATC GCG TTG GCC CAC CGC ATG TCG CGT CGA CTG ACC CGC GTG CGC      2461
Pro Ile Ala Leu Ala His Arg Met Ser Arg Arg Leu Thr Arg Val Arg
130                 135                 140                 145
```

```
AAC GAC GGC ACG CTG CCG TAC CTG CGT GCC GAC GGC AAG ACC CAG GTC      2509
Asn Asp Gly Thr Leu Pro Tyr Leu Arg Ala Asp Gly Lys Thr Gln Val
                150                 155                 160

ACC GTC GAG TAC GCC GGT GAC CAG CCG GTT CGC CTG GAC ACC ACG GTG      2557
Thr Val Glu Tyr Ala Gly Asp Gln Pro Val Arg Leu Asp Thr Thr Val
            165                 170                 175

CTG TCC AGC CAG CAC GCC GAG GAC GTC GAC CTC GAC AAG CAA CTG ATC      2605
Leu Ser Ser Gln His Ala Glu Asp Val Asp Leu Asp Lys Gln Leu Ile
        180                 185                 190

CCC GAG GTC AGG GAC AAG GTC ATC ACC CCG GAG ATC GAG AAG GTC GGG      2653
Pro Glu Val Arg Asp Lys Val Ile Thr Pro Glu Ile Glu Lys Val Gly
    195                 200                 205

CTG GAC ACC TCG GAC ATG CGT CTG CTG GTG AAT CCG ACG GGT CGG TTC      2701
Leu Asp Thr Ser Asp Met Arg Leu Leu Val Asn Pro Thr Gly Arg Phe
210                 215                 220                 225

GTC ACG GGT GGT CCG ATG GGG TGACTGCGGG CTGACCGGCC GCAAAGATCA         2752
Val Thr Gly Gly Pro Met Gly  *
                230
                 SAMS◄

TCGTCGACAC CTACGGCGGG ATGGCCCCGC CACGGTGGCG GTGCCTCTCC GGTAAAGGAC    2812

CT                                                                   2814
```

Fig. 10

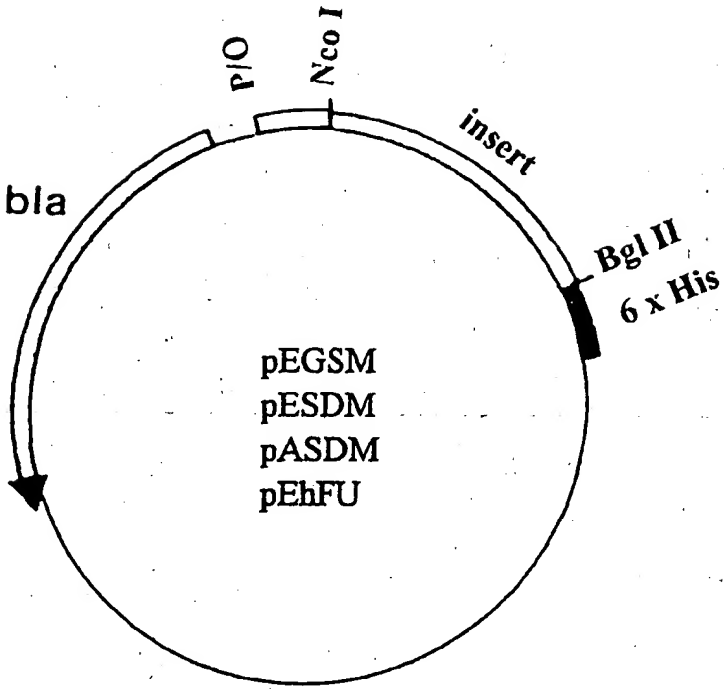


Fig. 11

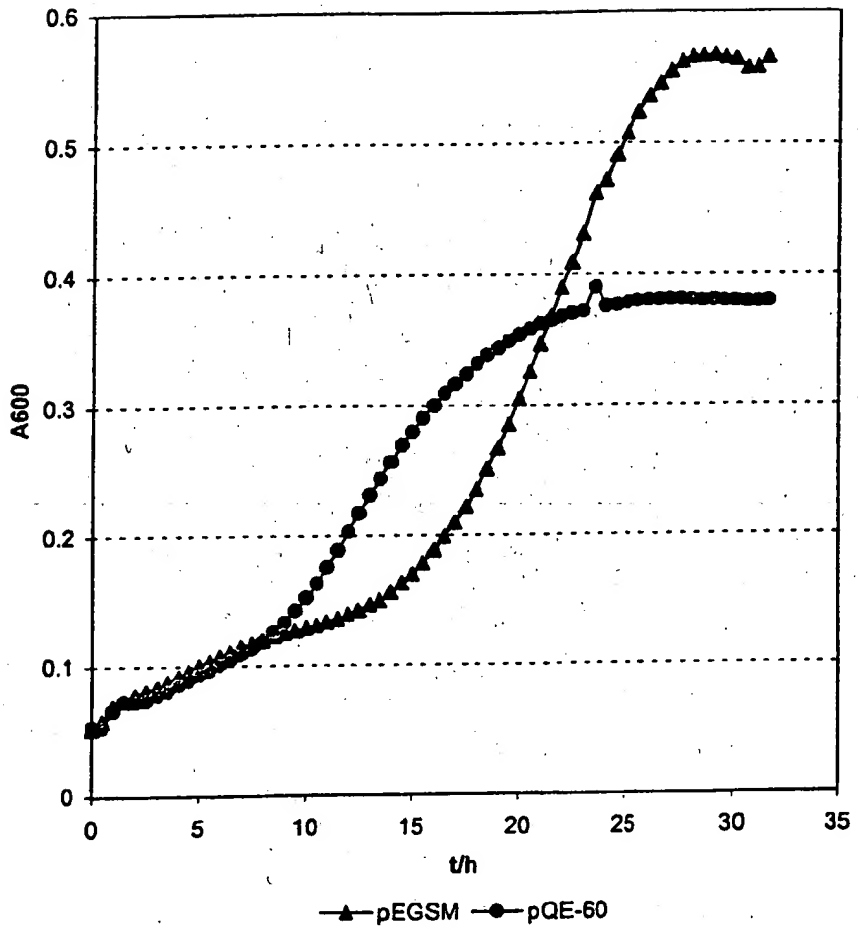
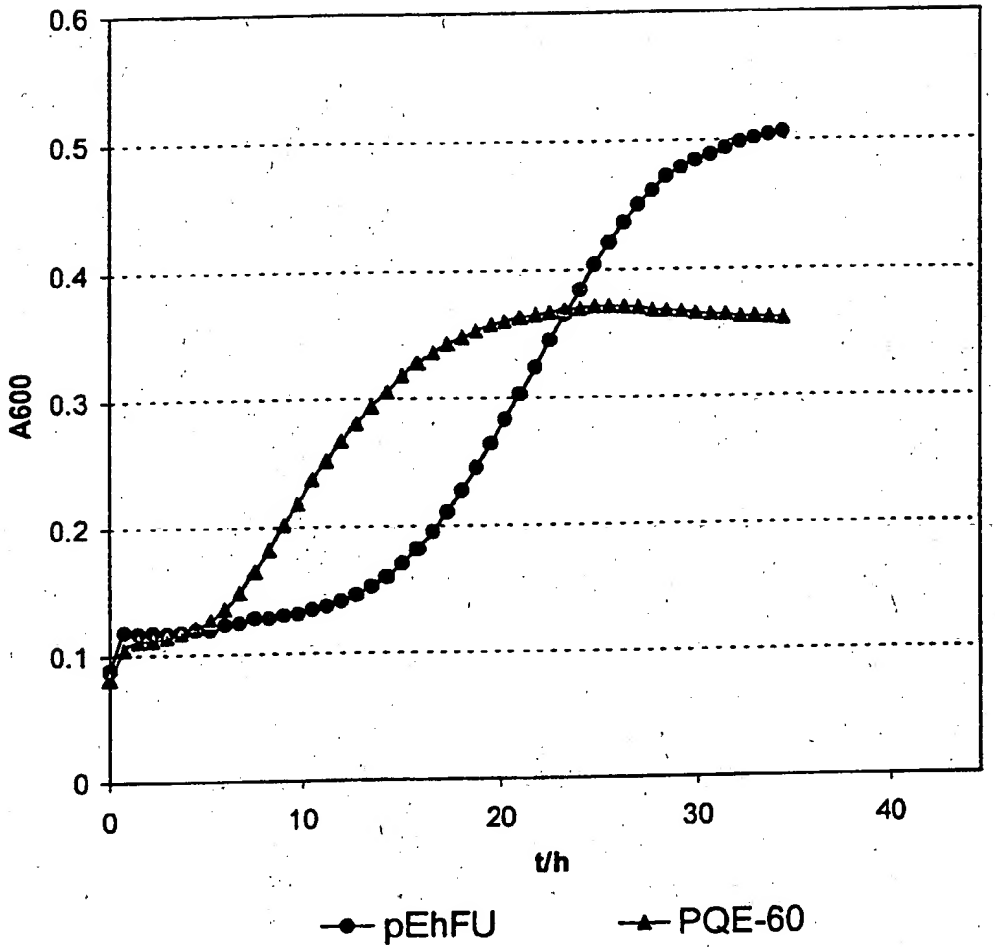


Fig. 12



SEQUENCE LISTING

<110> Reinikainen, Tapani
Nyyssölä, Antti
Kerovuo, Janne

<120> Methyltransferases, nucleic acid molecules encoding methyltransferases, their recombinant expression and uses thereof

<130> 73695

<140>

<141>

<160> 30

<170> PatentIn Ver. 2.0

<210> 1

<211> 2814

<212> DNA

<213> Actinopolyspora halophila

<220>

<221> CDS

<222> (208)..(1047)

<220>

<221> CDS

<222> (1048)..(1902)

<220>

<221> CDS

<222> (2027)..(2722)

<400> 1

aatctgcgcc gggatcggac ggcaaaggtg ccgaccgagt gtgctgatcg atgccagcgc 60

gcggcccgct agttcggctg cccgcatcgg tccggcgtgc cccttgtgtg aggtgacccg 120

gcgcgcccgt gaaagctcgg ggcggcgtaa ttcgccgtgg ttccgagcga tgcagtggcg 180

cttgaccaac gcgagtggga gctgagt atg acc aag agc gtg gac gat ctt gcc 234
Met Thr Lys Ser Val Asp Asp Leu Ala
1 5

cgt ggt gac cag gcc ggg gac gag cag gac ccg gtg cac cgc gag cag 282
Arg Gly Asp Gln Ala Gly Asp Glu Gln Asp Pro Val His Arg Glu Gln
10 15 20 25

cag acg ttc ggc gac aat ccg ttg gaa gta cgc gac act gat cac tac 330
Gln Thr Phe Gly Asp Asn Pro Leu Glu Val Arg Asp Thr Asp His Tyr
30 35 40

atg cat gag tac gtc ggt ggt ttt gtc gac aag tgg gac gat ctg atc 378
Met His Glu Tyr Val Gly Gly Phe Val Asp Lys Trp Asp Asp Leu Ile
45 50 55

gat tgg aag aag cgc tac gaa agc gag ggc agc ttc ttc atc gac caa 426
Asp Trp Lys Lys Arg Tyr Glu Ser Glu Gly Ser Phe Phe Ile Asp Gln
60 65 70

ttg cgc gca cgc ggt gtc gag acc gtg ctg gac gcg gcg gcc ggg acc 474
Leu Arg Ala Arg Gly Val Glu Thr Val Leu Asp Ala Ala Ala Gly Thr
75 80 85

ggt ttc cac tcg gtc cgg ttg ctc gag gag ggg ttt gag acc gtc agc 522
Gly Phe His Ser Val Arg Leu Leu Glu Glu Gly Phe Glu Thr Val Ser
90 95 100 105

gcg gac ggc agc ccg cag atg ctg gcc aag gcc ttc agt aac gga ctg 570
Ala Asp Gly Ser Pro Gln Met Leu Ala Lys Ala Phe Ser Asn Gly Leu
110 115 120

gcc tac aac ggt cac att ctg cgt gtg gtc aac gcg gac tgg cgt tgg 618
Ala Tyr Asn Gly His Ile Leu Arg Val Val Asn Ala Asp Trp Arg Trp
125 130 135

ctc aac cgt gac gtg cac ggt gaa tac gac gcg atc att tgc ctg ggc 666
Leu Asn Arg Asp Val His Gly Glu Tyr Asp Ala Ile Ile Cys Leu Gly
140 145 150

aac tcc ttt acc cac ctg ttc tcg gag cgg gac cgc cgc aag acg ctg 714
Asn Ser Phe Thr His Leu Phe Ser Glu Arg Asp Arg Arg Lys Thr Leu
155 160 165

gct gag ttc tac gcg atg ctc aag cac gac ggt gtc ctg atc atc gac 762
Ala Glu Phe Tyr Ala Met Leu Lys His Asp Gly Val Leu Ile Ile Asp
170 175 180 185

cag cga aac tac gac tcc att ctt gac acc ggc ttc tcc agt aag cac 810
Gln Arg Asn Tyr Asp Ser Ile Leu Asp Thr Gly Phe Ser Ser Lys His
190 195 200

acg tat tac tac gcc ggt gag gac gtt tcc gcg gag ccc gac cac atc 858
Thr Tyr Tyr Tyr Ala Gly Glu Asp Val Ser Ala Glu Pro Asp His Ile
205 210 215

gac gac ggg ctg gcg cgg ttc aag tac acg ttc ccg gac aag tcc gaa 906
Asp Asp Gly Leu Ala Arg Phe Lys Tyr Thr Phe Pro Asp Lys Ser Glu
220 225 230

ttc ttc ctg aac atg tac ccg ctg cgg aaa gac tac atg cgg cgg ctc 954

Phe Phe Leu Asn Met Tyr Pro Leu Arg Lys Asp Tyr Met Arg Arg Leu
235 240 245

atg cgt gag gtc ggt ttc caa agg att gac acc tac ggt gat ttc cag 1002
Met Arg Glu Val Gly Phe Gln Arg Ile Asp Thr Tyr Gly Asp Phe Gln
250 255 260 265

gaa act tac ggt gaa gac gag ccc gac ttc tac atc cac gtc gcg gag 1050
Glu Thr Tyr Gly Glu Asp Glu Pro Asp Phe Tyr Ile His Val Ala Glu
270 275 280 1

aag agc tac cgc acc gag gac gag ttc gtc gac atg tac tcg aac gcg 1098
Lys Ser Tyr Arg Thr Glu Asp Glu Phe Val Asp Met Tyr Ser Asn Ala
5 10 15

gtg cac acc gcg cgg gac tac tac aac tcc gag gac gcg gac aac ttc 1146
Val His Thr Ala Arg Asp Tyr Tyr Asn Ser Glu Asp Ala Asp Asn Phe
20 25 30

tac tac cac gtc tgg ggc ggc aac gac atc cac gtc ggg ctg tac cag 1194
Tyr Tyr His Val Trp Gly Gly Asn Asp Ile His Val Gly Leu Tyr Gln
35 40 45

aca ccg cag gag gac atc gcc acc gcc agt gag cgc act gtc cag cgg 1242
Thr Pro Gln Glu Asp Ile Ala Thr Ala Ser Glu Arg Thr Val Gln Arg
50 55 60 65

atg gcg ggc aag gtc gac atc agc ccc gaa acc agg att ctg gat ctc 1290
Met Ala Gly Lys Val Asp Ile Ser Pro Glu Thr Arg Ile Leu Asp Leu
70 75 80

ggt gcc ggc tac ggc gga gcc gcg cgg tac ctg gcc agg acc tac ggc 1338
Gly Ala Gly Tyr Gly Gly Ala Ala Arg Tyr Leu Ala Arg Thr Tyr Gly
85 90 95

tgc cac gtc acc tgc ctc aac ctc agc gag gtg gag aac cag cgc aac 1386
Cys His Val Thr Cys Leu Asn Leu Ser Glu Val Glu Asn Gln Arg Asn
100 105 110

cgc gag atc act cgc gcc gag ggg ctc gag cac ctg atc gag gtg acc 1434
Arg Glu Ile Thr Arg Ala Glu Gly Leu Glu His Leu Ile Glu Val Thr
115 120 125

gac ggt tcc ttc gag gat ctc ccc tac cag gac aac gcg ttc gac gtg 1482
Asp Gly Ser Phe Glu Asp Leu Pro Tyr Gln Asp Asn Ala Phe Asp Val
130 135 140 145

gtc tgg tcg cag gac tcc ttc ctc cac agc ggt gac cgc agc agg gtc 1530
Val Trp Ser Gln Asp Ser Phe Leu His Ser Gly Asp Arg Ser Arg Val
150 155 160

atg gaa gag gtg acc cgg gtc ctc aag ccg aag ggt tcg gtg ctg ttc 1578
Met Glu Glu Val Thr Arg Val Leu Lys Pro Lys Gly Ser Val Leu Phe
165 170 175

acc gat ccg atg gcg tcc gac tcg gcg aag aag aac gag ctc ggc ccc 1626
Thr Asp Pro Met Ala Ser Asp Ser Ala Lys Lys Asn Glu Leu Gly Pro
180 185 190

atc ctg gac agg ctg cac ctg gac tcg ctc ggc tcg ccc ggt ttc tac 1674
Ile Leu Asp Arg Leu His Leu Asp Ser Leu Gly Ser Pro Gly Phe Tyr
195 200 205

cgg aag gag ctg act cgt ctc ggg ctg cag aac atc gag ttc gag gac 1722
Arg Lys Glu Leu Thr Arg Leu Gly Leu Gln Asn Ile Glu Phe Glu Asp
210 215 220 225

ctc agc gaa tac ctg ccc gtc cac tac ggc cgg gtt ctg gaa gtg ctg 1770

Leu Ser Glu Tyr Leu Pro Val His Tyr Gly Arg Val Leu Glu Val Leu
230 235 240

gag agc cgg gag aac gag ctc gcc ggc ttc atc ggc gag gag tac cga 1818
Glu Ser Arg Glu Asn Glu Leu Ala Gly Phe Ile Gly Glu Glu Tyr Arg
245 250 255

gct cac atg aag acc ggg ctg cgc aac tgg gtg cag gcc ggc aat ggc 1866
Ala His Met Lys Thr Gly Leu Arg Asn Trp Val Gln Ala Gly Asn Gly
260 265 270

ggg agc ctg gcc tgg ggc atc atc cac gcc agg gca tgacagccgt 1912
Gly Ser Leu Ala Trp Gly Ile Ile His Ala Arg Ala
275 280 285

cgcgtacgaa gaacgatcag aagcagttag cagtgaggtg aagatcagcc gtgactgaga 1972

tgaaccgcag gttgttcacc aagtgagtcc gtgaccgagg gccacccgga caag atg 2029
Met
1

gcc gac tcg atc agc gac gcg atc ctg gac gcg atg ctg gct cag gac 2077
Ala Asp Ser Ile Ser Asp Ala Ile Leu Asp Ala Met Leu Ala Gln Asp
5 10 15

ccc cgc tcc cgc gtg gcc atg gag acc atg atc acc acc ggg cag gtg 2125
Pro Arg Ser Arg Val Ala Met Glu Thr Met Ile Thr Thr Gly Gln Val
20 25 30

cac ctg gcc ggt gag gtg acc acc gag gcc gac gtc gac ctg ccc gcg 2173
His Leu Ala Gly Glu Val Thr Thr Glu Ala Asp Val Asp Leu Pro Ala
35 40 45

atc gtg cgg gag aag gtc ctc gag atc ggc tac gac aac tcg gcc aag 2221

Ile Val Arg Glu Lys Val Leu Glu Ile Gly Tyr Asp Asn Ser Ala Lys
50 55 60 65

ggc ttc gac gga gac tcc tgc ggc atc aac gtc tcc atc gac gcg cag 2269
Gly Phe Asp Gly Asp Ser Cys Gly Ile Asn Val Ser Ile Asp Ala Gln
70 75 80

tcc ccg gac atc ggc cag ggc gtg gac tcc gct cac gag tcc cgc gtc 2317
Ser Pro Asp Ile Gly Gln Gly Val Asp Ser Ala His Glu Ser Arg Val
85 90 95

gag ggt gcc atc gac gag atc gcc agt cag ggc gcc ggc gac cag ggc 2365
Glu Gly Ala Ile Asp Glu Ile Ala Ser Gln Gly Ala Gly Asp Gln Gly
100 105 110

ctg atg ttc ggt tac gcc acc agc gag acc gac gag ctc atg ccg ctg 2413
Leu Met Phe Gly Tyr Ala Thr Ser Glu Thr Asp Glu Leu Met Pro Leu
115 120 125

ccg atc gcg ttg gcc cac cgc atg tcg cgt cga ctg acc cgc gtg cgc 2461
Pro Ile Ala Leu Ala His Arg Met Ser Arg Arg Leu Thr Arg Val Arg
130 135 140 145

aac gac ggc acg ctg ccg tac ctg cgt gcc gac ggc aag acc cag gtc 2509
Asn Asp Gly Thr Leu Pro Tyr Leu Arg Ala Asp Gly Lys Thr Gln Val
150 155 160

acc gtc gag tac gcc ggt gac cag ccg gtt cgc ctg gac acc acg gtg 2557
Thr Val Glu Tyr Ala Gly Asp Gln Pro Val Arg Leu Asp Thr Thr Val
165 170 175

ctg tcc agc cag cac gcc gag gac gtc gac ctc gac aag caa ctg atc 2605
Leu Ser Ser Gln His Ala Glu Asp Val Asp Leu Asp Lys Gln Leu Ile
180 185 190

ccc gag gtc agg gac aag gtc atc acc ccg gag atc gag aag gtc ggg 2653
Pro Glu Val Arg Asp Lys Val Ile Thr Pro Glu Ile Glu Lys Val Gly
195 200 205

ctg gac acc tcg gac atg cgt ctg ctg gtg aat ccg acg ggt cgg ttc 2701
Leu Asp Thr Ser Asp Met Arg Leu Leu Val Asn Pro Thr Gly Arg Phe
210 215 220 225

gtc acg ggt ggt ccg atg ggg tgactgcggg ctgaccggcc gcaaagatca 2752
Val Thr Gly Gly Pro Met Gly
230

tcgtcgacac ctacggcggg atggccccgc cacggtggcg gtgcctctcc ggtaaaggac 2812

ct 2814

<210> 2
<211> 280
<212> PRT
<213> Actinopolyspora halophila

<400> 2
Met Thr Lys Ser Val Asp Asp Leu Ala Arg Gly Asp Gln Ala Gly Asp
1 5 10 15

Glu Gln Asp Pro Val His Arg Glu Gln Gln Thr Phe Gly Asp Asn Pro
20 25 30

Leu Glu Val Arg Asp Thr Asp His Tyr Met His Glu Tyr Val Gly Gly
35 40 45

Phe Val Asp Lys Trp Asp Asp Leu Ile Asp Trp Lys Lys Arg Tyr Glu

50 55 60

Ser Glu Gly Ser Phe Phe Ile Asp Gln Leu Arg Ala Arg Gly Val Glu
65 70 75 80

Thr Val Leu Asp Ala Ala Ala Gly Thr Gly Phe His Ser Val Arg Leu
85 90 95

Leu Glu Glu Gly Phe Glu Thr Val Ser Ala Asp Gly Ser Pro Gln Met
100 105 110

Leu Ala Lys Ala Phe Ser Asn Gly Leu Ala Tyr Asn Gly His Ile Leu
115 120 125

Arg Val Val Asn Ala Asp Trp Arg Trp Leu Asn Arg Asp Val His Gly
130 135 140

Glu Tyr Asp Ala Ile Ile Cys Leu Gly Asn Ser Phe Thr His Leu Phe
145 150 155 160

Ser Glu Arg Asp Arg Arg Lys Thr Leu Ala Glu Phe Tyr Ala Met Leu
165 170 175

Lys His Asp Gly Val Leu Ile Ile Asp Gln Arg Asn Tyr Asp Ser Ile
180 185 190

Leu Asp Thr Gly Phe Ser Ser Lys His Thr Tyr Tyr Tyr Ala Gly Glu
195 200 205

Asp Val Ser Ala Glu Pro Asp His Ile Asp Asp Gly Leu Ala Arg Phe
210 215 220

Lys Tyr Thr Phe Pro Asp Lys Ser Glu Phe Phe Leu Asn Met Tyr Pro
225 230 235 240

Leu Arg Lys Asp Tyr Met Arg Arg Leu Met Arg Glu Val Gly Phe Gln
245 250 255

Arg Ile Asp Thr Tyr Gly Asp Phe Gln Glu Thr Tyr Gly Glu Asp Glu
260 265 270

Pro Asp Phe Tyr Ile His Val Ala
275 280

<210> 3
<211> 285
<212> PRT
<213> Actinopolyspora halophila

<400> 3
Glu Lys Ser Tyr Arg Thr Glu Asp Glu Phe Val Asp Met Tyr Ser Asn
1 5 10 15

Ala Val His Thr Ala Arg Asp Tyr Tyr Asn Ser Glu Asp Ala Asp Asn
20 25 30

Phe Tyr Tyr His Val Trp Gly Gly Asn Asp Ile His Val Gly Leu Tyr
35 40 45

Gln Thr Pro Gln Glu Asp Ile Ala Thr Ala Ser Glu Arg Thr Val Gln
50 55 60

Arg Met Ala Gly Lys Val Asp Ile Ser Pro Glu Thr Arg Ile Leu Asp
65 70 75 80

L u Gly Ala Gly Tyr Gly Gly Ala Ala Arg Tyr Leu Ala Arg Thr Tyr
85 90 95

Gly Cys His Val Thr Cys Leu Asn Leu Ser Glu Val Glu Asn Gln Arg
100 105 110

Asn Arg Glu Ile Thr Arg Ala Glu Gly Leu Glu His Leu Ile Glu Val
115 120 125

Thr Asp Gly Ser Phe Glu Asp Leu Pro Tyr Gln Asp Asn Ala Phe Asp
130 135 140

Val Val Trp Ser Gln Asp Ser Phe Leu His Ser Gly Asp Arg Ser Arg
145 150 155 160

Val Met Glu Glu Val Thr Arg Val Leu Lys Pro Lys Gly Ser Val Leu
165 170 175

Phe Thr Asp Pro Met Ala Ser Asp Ser Ala Lys Lys Asn Glu Leu Gly
180 185 190

Pro Ile Leu Asp Arg Leu His Leu Asp Ser Leu Gly Ser Pro Gly Phe
195 200 205

Tyr Arg Lys Glu Leu Thr Arg Leu Gly Leu Gln Asn Ile Glu Phe Glu
210 215 220

Asp Leu Ser Glu Tyr Leu Pro Val His Tyr Gly Arg Val Leu Glu Val
225 230 235 240

Leu Glu Ser Arg Glu Asn Glu Leu Ala Gly Phe Ile Gly Glu Glu Tyr
245 250 255

Arg Ala His Met Lys Thr Gly Leu Arg Asn Trp Val Gln Ala Gly Asn
260 265 270

Gly Gly Ser Leu Ala Trp Gly Ile Ile His Ala Arg Ala
275 280 285

<210> 4
<211> 232
<212> PRT
<213> Actinopolyspora halophila

<400> 4
Met Ala Asp Ser Ile Ser Asp Ala Ile Leu Asp Ala Met Leu Ala Gln
1 5 10 15

Asp Pro Arg Ser Arg Val Ala Met Glu Thr Met Ile Thr Thr Gly Gln
20 25 30

Val His Leu Ala Gly Glu Val Thr Thr Glu Ala Asp Val Asp Leu Pro
35 40 45

Ala Ile Val Arg Glu Lys Val Leu Glu Ile Gly Tyr Asp Asn Ser Ala
50 55 60

Lys Gly Phe Asp Gly Asp Ser Cys Gly Ile Asn Val Ser Ile Asp Ala
65 70 75 80

Gln Ser Pro Asp Ile Gly Gln Gly Val Asp Ser Ala His Glu Ser Arg
85 90 95

Val Glu Gly Ala Ile Asp Glu Ile Ala Ser Gln Gly Ala Gly Asp Gln
100 105 110

Gly Leu Met Phe Gly Tyr Ala Thr Ser Glu Thr Asp Glu Leu Met Pro
115 120 125

Leu Pro Ile Ala Leu Ala His Arg Met Ser Arg Arg Leu Thr Arg Val
130 135 140

Arg Asn Asp Gly Thr Leu Pro Tyr Leu Arg Ala Asp Gly Lys Thr Gln
145 150 155 160

Val Thr Val Glu Tyr Ala Gly Asp Gln Pro Val Arg Leu Asp Thr Thr
165 170 175

Val Leu Ser Ser Gln His Ala Glu Asp Val Asp Leu Asp Lys Gln Leu
180 185 190

Ile Pro Glu Val Arg Asp Lys Val Ile Thr Pro Glu Ile Glu Lys Val
195 200 205

Gly Leu Asp Thr Ser Asp Met Arg Leu Leu Val Asn Pro Thr Gly Arg
210 215 220

Phe Val Thr Gly Gly Pro Met Gly
225 230

<210> 5
<211> 3219
<212> DNA
<213> Ectothiorhodospira halochloris

<220>
<221> CDS
<222> (221)..(1024)

<220>
<221> CDS
<222> (1031)..(1867)

<220>
<221> CDS
<222> (2006)..(3004)

<400> 5
gcaggctcgg tgccaccaat acagggcagc cccctggcca gaaacacgcc ccgtagttac 60

ggaggcgcat gccatagata atgaacttgc tacaagaggc acttgacatt cagttgtcca 120

agcggcgccg gcgatcaagc acacgatctg aattcacagc gtatgggtat accgtagcgc 180

gcgttcgcta caccaaccat tcgccaagct gaggtgatat atg aat acg act act 235
Met Asn Thr Thr Thr
1 5

gag cag gac ttc gga gcg gac cca aca aaa gtc cga gat acc gac cac 283
Glu Gln Asp Phe Gly Ala Asp Pro Thr Lys Val Arg Asp Thr Asp His
10 15 20

tac acg gaa gag tat gtt gac ggg ttt gtc gac aaa tgg gac gac ttg 331
Tyr Thr Glu Glu Tyr Val Asp Gly Phe Val Asp Lys Trp Asp Asp Leu
25 30 35

att gat tgg gat agc cgc gcg aag agt gaa ggt gat ttc ttt att caa 379
Ile Asp Trp Asp Ser Arg Ala Lys Ser Glu Gly Asp Phe Phe Ile Gln
40 45 50

gag ttg aag aag cgt ggt gcc aca cgc atc ctc gat gcc gca aca ggt 427
Glu Leu Lys Lys Arg Gly Ala Thr Arg Ile Leu Asp Ala Ala Thr Gly
55 60 65

acc ggc ttc cac tca gtg cgt ttg ctc gag gcg ggt ttt gat gtc gtc 475
Thr Gly Phe His Ser Val Arg Leu Leu Glu Ala Gly Phe Asp Val Val

70 75 80 85

agt gct gat ggc agt gcc gag atg ctc gcc aaa gcc ttc gag aat ggg 523
Ser Ala Asp Gly Ser Ala Glu Met Leu Ala Lys Ala Phe Glu Asn Gly
90 95 100

cgt aag cgt ggc cat atc cta cgc acg gtt cag gtc gac tgg cgg tgg 571
Arg Lys Arg Gly His Ile Leu Arg Thr Val Gln Val Asp Trp Arg Trp
105 110 115

ctg aat cgc gat atc cac ggt cgt tat gac gcc att atc tgc ctt ggc 619
Leu Asn Arg Asp Ile His Gly Arg Tyr Asp Ala Ile Ile Cys Leu Gly
120 125 130

aac tcg ttt acc cac ctg ttt aat gaa aaa gat cgg cgc aag acc ctg 667
Asn Ser Phe Thr His Leu Phe Asn Glu Lys Asp Arg Arg Lys Thr Leu
135 140 145

gca gag ttc tat tcc gcg ctg aac ccg gaa ggg gtg ctg ata ttg gat 715
Ala Glu Phe Tyr Ser Ala Leu Asn Pro Glu Gly Val Leu Ile Leu Asp
150 155 160 165

cag cgc aat tac gac ggc atc ctc gat cac ggc tat gac tca agc cac 763
Gln Arg Asn Tyr Asp Gly Ile Leu Asp His Gly Tyr Asp Ser Ser His
170 175 180

tcc tac tac tat tgc ggt gag ggc gtc tct gtc tat cct gag cac gtt 811
Ser Tyr Tyr Tyr Cys Gly Glu Gly Val Ser Val Tyr Pro Glu His Val
185 190 195

gat gat ggc cta gcc cgc ttc aag tat gaa ttc aac gac ggt tca act 859
Asp Asp Gly Leu Ala Arg Phe Lys Tyr Glu Phe Asn Asp Gly Ser Thr
200 205 210

```
tac ttc ctg aac atg ttc ccg ctg cgc aag gat tac acg cgc cgc cta   907
Tyr Phe Leu Asn Met Phe Pro Leu Arg Lys Asp Tyr Thr Arg Arg Leu
    215                 220                 225

atg cat gag gtg ggt ttc caa aaa atc gat acc tac ggc gac ttc aag   955
Met His Glu Val Gly Phe Gln Lys Ile Asp Thr Tyr Gly Asp Phe Lys
230                 235                 240                 245

gca acc tac cgc gat gca gat ccg gat ttc ttt att cat gtc gcc gag   1003
Ala Thr Tyr Arg Asp Ala Asp Pro Asp Phe Phe Ile His Val Ala Glu
                250                 255                 260

aag gaa tat cgg gag gag gac tgatat atg gcg acg cgc tac gac gat   1051
Lys Glu Tyr Arg Glu Glu Asp        Met Ala Thr Arg Tyr Asp Asp
            265                     1               5

caa gcc ata gaa acg gcg agg cag tac tat aac agt gag gat gcc gat   1099
Gln Ala Ile Glu Thr Ala Arg Gln Tyr Tyr Asn Ser Glu Asp Ala Asp
        10                  15                  20

aac ttc tac gcc ata ata tgg ggt ggt gag gat atc cat atc ggt ttg   1147
Asn Phe Tyr Ala Ile Ile Trp Gly Gly Glu Asp Ile His Ile Gly Leu
    25                  30                  35

tac aac gat gat gaa gag ccg ata gca gat gcc agc agg cgt act gta   1195
Tyr Asn Asp Asp Glu Glu Pro Ile Ala Asp Ala Ser Arg Arg Thr Val
40                  45                  50                  55

gag cgc atg tcg tca ctg tcc cgg cag ctt ggc cca gat agc tat gtc   1243
Glu Arg Met Ser Ser Leu Ser Arg Gln Leu Gly Pro Asp Ser Tyr Val
                60                  65                  70

ctc gac atg ggg gct ggt tac gga ggg tct gcc cgt tac ctt gcg cat   1291
Leu Asp Met Gly Ala Gly Tyr Gly Gly Ser Ala Arg Tyr Leu Ala His
```

75 80 85

aag tat ggc tgt aag gtg gcg gcg ctc aat ctc tcc gaa cgt gag aat 1339
Lys Tyr Gly Cys Lys Val Ala Ala Leu Asn Leu Ser Glu Arg Glu Asn
90 95 100

gag cgt gat cgg cag atg aat aag gag caa gga gtc gat cac ctg att 1387
Glu Arg Asp Arg Gln Met Asn Lys Glu Gln Gly Val Asp His Leu Ile
105 110 115

gag gtt gtc gat gct gcc ttt gaa gat gtc ccc tat gat gat ggg gtg 1435
Glu Val Val Asp Ala Ala Phe Glu Asp Val Pro Tyr Asp Asp Gly Val
120 125 130 135

ttt gac ctg gtc tgg tcg cag gac tcc ttc ctc cat agt ccg gat agg 1483
Phe Asp Leu Val Trp Ser Gln Asp Ser Phe Leu His Ser Pro Asp Arg
140 145 150

gaa cgc gta ctg cgg gag gct agt aga gtt ctg cgt tct ggg ggc gag 1531
Glu Arg Val Leu Arg Glu Ala Ser Arg Val Leu Arg Ser Gly Gly Glu
155 160 165

ttt atc ttc acc gat cca atg cag gcg gat gat tgc ccc gag ggg gtt 1579
Phe Ile Phe Thr Asp Pro Met Gln Ala Asp Asp Cys Pro Glu Gly Val
170 175 180

att caa ccg atc ctc gat agg atc cac ctc gag acg atg ggc acg cct 1627
Ile Gln Pro Ile Leu Asp Arg Ile His Leu Glu Thr Met Gly Thr Pro
186 190 195

aat ttc tat cgc cag acc ctt agg gat ctg ggc ttt gag gaa atc acc 1675
Asn Phe Tyr Arg Gln Thr Leu Arg Asp Leu Gly Phe Glu Glu Ile Thr
200 205 210 215

ttc gaa gac cac acc cat cag ttg ccg cgc cat tac ggg cgc gtg cgc 1723
Phe Glu Asp His Thr His Gln Leu Pro Arg His Tyr Gly Arg Val Arg
220 225 230

cgt gag cta gat cgt cga gaa ggt gaa ctg cag ggg cat gtc tcg gca 1771
Arg Glu Leu Asp Arg Arg Glu Gly Glu Leu Gln Gly His Val Ser Ala
235 240 245

gag tac atc gaa cgc atg aaa aat ggc ctt gac cac tgg gtc aat ggc 1819
Glu Tyr Ile Glu Arg Met Lys Asn Gly Leu Asp His Trp Val Asn Gly
250 255 260

ggt aac aag ggt tat ctc acc tgg ggt atc ttt tac ttc cgc aaa ggg 1867
Gly Asn Lys Gly Tyr Leu Thr Trp Gly Ile Phe Tyr Phe Arg Lys Gly
265 270 275

tgatcactag agcgctaaat cgcaggcgcg gtaattgtgc cgcgcttcgg gcgcatatgt 1927

tttagcccag tggatttgca catagaaaag agtgaattct ggtttagcgc tcgaataaac 1987

tatagcaaag gagaaact atg act aag cga tat cta ttt acc tct gag tcg 2038
Met Thr Lys Arg Tyr Leu Phe Thr Ser Glu Ser
1 5 10

gtc tct gaa ggc cac ccg gac aaa atg gcc gac cag att tcg gat gca 2086
Val Ser Glu Gly His Pro Asp Lys Met Ala Asp Gln Ile Ser Asp Ala
15 20 25

ctg ctc gat gag ttc ctg cgc cag gat cca aag tcg cgg gtg gct gca 2134
Leu Leu Asp Glu Phe Leu Arg Gln Asp Pro Lys Ser Arg Val Ala Ala
30 35 40

gag acg atg att cag acc ggc atg gtt gta gtt gcc ggc gaa att aag 2182
Glu Thr Met Ile Gln Thr Gly Met Val Val Val Ala Gly Glu Ile Lys

45 50 55

agc aat gcc aag atc aat gtt gag ccg ctg gta cgt gaa gtt gtc cgc 2230
Ser Asn Ala Lys Ile Asn Val Glu Pro Leu Val Arg Glu Val Val Arg
60 65 70 75

gat atc ggc tac acc agc tca gat atg ggc ttt gat gcc gac acc tgt 2278
Asp Ile Gly Tyr Thr Ser Ser Asp Met Gly Phe Asp Ala Asp Thr Cys
80 85 90

gcc gta ctc aac gcc ctc ggc gag cag tcc ccc gac atc aat caa ggc 2326
Ala Val Leu Asn Ala Leu Gly Glu Gln Ser Pro Asp Ile Asn Gln Gly
95 100 105

gtt gac cgg gaa gag gaa gag gag cag ggc gcc ggt gac caa gga ctg 2374
Val Asp Arg Glu Glu Glu Glu Glu Gln Gly Ala Gly Asp Gln Gly Leu
110 115 120

atg ttc ggt tac gcc acc aat gag acc gac gtc ctc atg ccg gca gcg 2422
Met Phe Gly Tyr Ala Thr Asn Glu Thr Asp Val Leu Met Pro Ala Ala
125 130 135

atc cac tac tcg cac ctg ctg gtt aag cgc caa tcc gaa gtc cgt aac 2470
Ile His Tyr Ser His Leu Leu Val Lys Arg Gln Ser Glu Val Arg Asn
140 145 150 155

tcc aag aag ctg ccg tgg ttg cgc ccg gac gct aag agc cag gtc acc 2518
Ser Lys Lys Leu Pro Trp Leu Arg Pro Asp Ala Lys Ser Gln Val Thr
160 165 170

ttc aag tac gaa ggc gat aag atc gtc ggt tgc gat gcg gtg gtg ctc 2566
Phe Lys Tyr Glu Gly Asp Lys Ile Val Gly Cys Asp Ala Val Val Leu
175 180 185

```
tcc acg cag cat gac gag acc gtt gat cag aag acc gtc cac gag ggc   2614
Ser Thr Gln His Asp Glu Thr Val Asp Gln Lys Thr Val His Glu Gly
        190                 195                 200

gtg atg gaa gag atc atc aag ccc atc ctc ggt gac acc ggc tgg ctg   2662
Val Met Glu Glu Ile Ile Lys Pro Ile Leu Gly Asp Thr Gly Trp Leu
    205                 210                 215

acc aac gag acc aag tac cac att aac ccg acc ggc cgc ttc gtt act   2710
Thr Asn Glu Thr Lys Tyr His Ile Asn Pro Thr Gly Arg Phe Val Thr
220                 225                 230                 235

ggt ggt ccg ttg ggt gac tgt ggc ttg acc ggc cgt aag atc atc gtc   2758
Gly Gly Pro Leu Gly Asp Cys Gly Leu Thr Gly Arg Lys Ile Ile Val
                240                 245                 250

gat acc tac ggc ggt atg ggc cgt cac ggc ggc ggt gcc ttc tcc ggc   2806
Asp Thr Tyr Gly Gly Met Gly Arg His Gly Gly Gly Ala Phe Ser Gly
            255                 260                 265

aaa gat ccg tct aag gtc gat cgc tcg gcg gct tat gtt ggc cgt tat   2854
Lys Asp Pro Ser Lys Val Asp Arg Ser Ala Ala Tyr Val Gly Arg Tyr
        270                 275                 280

gta gct aag aac ata gtg gct gcc ggt ctg gcc gac cgc tgt gag gtg   2902
Val Ala Lys Asn Ile Val Ala Ala Gly Leu Ala Asp Arg Cys Glu Val
    285                 290                 295

cag ctc tcc tac gcc atc ggt gtt gcc gag ccg acc tcg gtc aat gtc   2950
Gln Leu Ser Tyr Ala Ile Gly Val Ala Glu Pro Thr Ser Val Asn Val
300                 305                 310                 315

gaa acc ttt ggt acc ggc aag gtc gaa gaa gag ctc gcg cac gat ggc   2998
Glu Thr Phe Gly Thr Gly Lys Val Glu Glu Glu Leu Ala His Asp Gly
```

320 325 330

cag ccc taagccgtgg ccgctgcgcg agccctggta gccggcgcga tcgaacgccg 3054
Gln Pro

caaagatatg tgcctgctgc tctgggggga tccccggacc ggtatcgcgc acttggaagg 3114

tgattcggtc atcgagctcc accgcggtgg cggccgctct agaactagtg gatccccccgg 3174

gctgcaggaa ttcgatatca agcttatcga taccgtcgac ctcga 3219

<210> 6
<211> 268
<212> PRT
<213> Ectothiorhodospira halochloris

<400> 6
Met Asn Thr Thr Thr Glu Gln Asp Phe Gly Ala Asp Pro Thr Lys Val
1 5 10 15

Arg Asp Thr Asp His Tyr Thr Glu Glu Tyr Val Asp Gly Phe Val Asp
20 25 30

Lys Trp Asp Asp Leu Ile Asp Trp Asp Ser Arg Ala Lys Ser Glu Gly
35 40 45

Asp Phe Phe Ile Gln Glu Leu Lys Lys Arg Gly Ala Thr Arg Ile Leu
50 55 60

Asp Ala Ala Thr Gly Thr Gly Phe His Ser Val Arg Leu Leu Glu Ala
65 70 75 80

Gly Phe Asp Val Val Ser Ala Asp Gly Ser Ala Glu Met Leu Ala Lys
85 90 95

Ala Phe Glu Asn Gly Arg Lys Arg Gly His Ile Leu Arg Thr Val Gln
100 105 110

Val Asp Trp Arg Trp Leu Asn Arg Asp Ile His Gly Arg Tyr Asp Ala
115 120 125

Ile Ile Cys Leu Gly Asn Ser Phe Thr His Leu Phe Asn Glu Lys Asp
130 135 140

Arg Arg Lys Thr Leu Ala Glu Phe Tyr Ser Ala Leu Asn Pro Glu Gly
145 150 155 160

Val Leu Ile Leu Asp Gln Arg Asn Tyr Asp Gly Ile Leu Asp His Gly
165 170 175

Tyr Asp Ser Ser His Ser Tyr Tyr Tyr Cys Gly Glu Gly Val Ser Val
180 185 190

Tyr Pro Glu His Val Asp Asp Gly Leu Ala Arg Phe Lys Tyr Glu Phe
195 200 205

Asn Asp Gly Ser Thr Tyr Phe Leu Asn Met Phe Pro Leu Arg Lys Asp
210 215 220

Tyr Thr Arg Arg Leu Met His Glu Val Gly Phe Gln Lys Ile Asp Thr
225 230 235 240

Tyr Gly Asp Phe Lys Ala Thr Tyr Arg Asp Ala Asp Pro Asp Phe Phe
245 250 255

Ile His Val Ala Glu Lys Glu Tyr Arg Glu Glu Asp

260 265

<210> 7

<211> 279

<212> PRT

<213> Ectothiorhodospira halochloris

<400> 7

Met Ala Thr Arg Tyr Asp Asp Gln Ala Ile Glu Thr Ala Arg Gln Tyr
1 5 10 15

Tyr Asn Ser Glu Asp Ala Asp Asn Phe Tyr Ala Ile Ile Trp Gly Gly
20 25 30

Glu Asp Ile His Ile Gly Leu Tyr Asn Asp Asp Glu Glu Pro Ile Ala
35 40 45

Asp Ala Ser Arg Arg Thr Val Glu Arg Met Ser Ser Leu Ser Arg Gln
50 55 60

Leu Gly Pro Asp Ser Tyr Val Leu Asp Met Gly Ala Gly Tyr Gly Gly
65 70 75 80

Ser Ala Arg Tyr Leu Ala His Lys Tyr Gly Cys Lys Val Ala Ala Leu
85 90 95

Asn Leu Ser Glu Arg Glu Asn Glu Arg Asp Arg Gln Met Asn Lys Glu
100 105 110

Gln Gly Val Asp His Leu Ile Glu Val Val Asp Ala Ala Phe Glu Asp
115 120 125

Val Pro Tyr Asp Asp Gly Val Phe Asp Leu Val Trp Ser Gln Asp Ser

130 135 140

Phe Leu His Ser Pro Asp Arg Glu Arg Val Leu Arg Glu Ala Ser Arg
145 150 155 160

Val Leu Arg Ser Gly Gly Glu Phe Ile Phe Thr Asp Pro Met Gln Ala
165 170 175

Asp Asp Cys Pro Glu Gly Val Ile Gln Pro Ile Leu Asp Arg Ile His
180 185 190

Leu Glu Thr Met Gly Thr Pro Asn Phe Tyr Arg Gln Thr Leu Arg Asp
195 200 205

Leu Gly Phe Glu Glu Ile Thr Phe Glu Asp His Thr His Gln Leu Pro
210 215 220

Arg His Tyr Gly Arg Val Arg Arg Glu Leu Asp Arg Arg Glu Gly Glu
225 230 235 240

Leu Gln Gly His Val Ser Ala Glu Tyr Ile Glu Arg Met Lys Asn Gly
245 250 255

Leu Asp His Trp Val Asn Gly Gly Asn Lys Gly Tyr Leu Thr Trp Gly
260 265 270

Ile Phe Tyr Phe Arg Lys Gly
275

<210> 8
<211> 333
<212> PRT
<213> Ectothiorhodospira halochloris

<400> 8

Met Thr Lys Arg Tyr Leu Phe Thr Ser Glu Ser Val Ser Glu Gly His
1 5 10 15

Pro Asp Lys Met Ala Asp Gln Ile Ser Asp Ala Leu Leu Asp Glu Phe
20 25 30

Leu Arg Gln Asp Pro Lys Ser Arg Val Ala Ala Glu Thr Met Ile Gln
35 40 45

Thr Gly Met Val Val Val Ala Gly Glu Ile Lys Ser Asn Ala Lys Ile
50 55 60

Asn Val Glu Pro Leu Val Arg Glu Val Val Arg Asp Ile Gly Tyr Thr
65 70 75 80

Ser Ser Asp Met Gly Phe Asp Ala Asp Thr Cys Ala Val Leu Asn Ala
85 90 95

Leu Gly Glu Gln Ser Pro Asp Ile Asn Gln Gly Val Asp Arg Glu Glu
100 105 110

Glu Glu Glu Gln Gly Ala Gly Asp Gln Gly Leu Met Phe Gly Tyr Ala
115 120 125

Thr Asn Glu Thr Asp Val Leu Met Pro Ala Ala Ile His Tyr Ser His
130 135 140

Leu Leu Val Lys Arg Gln Ser Glu Val Arg Asn Ser Lys Lys Leu Pro
145 150 155 160

Trp Leu Arg Pro Asp Ala Lys Ser Gln Val Thr Phe Lys Tyr Glu Gly
165 170 175

Asp Lys Ile Val Gly Cys Asp Ala Val Val Leu Ser Thr Gln His Asp
180 185 190

Glu Thr Val Asp Gln Lys Thr Val His Glu Gly Val Met Glu Glu Ile
195 200 205

Ile Lys Pro Ile Leu Gly Asp Thr Gly Trp Leu Thr Asn Glu Thr Lys
210 215 220

Tyr His Ile Asn Pro Thr Gly Arg Phe Val Thr Gly Gly Pro Leu Gly
225 230 235 240

Asp Cys Gly Leu Thr Gly Arg Lys Ile Ile Val Asp Thr Tyr Gly Gly
245 250 255

Met Gly Arg His Gly Gly Gly Ala Phe Ser Gly Lys Asp Pro Ser Lys
260 265 270

Val Asp Arg Ser Ala Ala Tyr Val Gly Arg Tyr Val Ala Lys Asn Ile
275 280 285

Val Ala Ala Gly Leu Ala Asp Arg Cys Glu Val Gln Leu Ser Tyr Ala
290 295 300

Ile Gly Val Ala Glu Pro Thr Ser Val Asn Val Glu Thr Phe Gly Thr
305 310 315 320

Gly Lys Val Glu Glu Glu Leu Ala His Asp Gly Gln Pro
325 330

<210> 9

<211> 37

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide

<400> 9

Glu Lys Ser Tyr Arg Thr Glu Asp Glu Phe Val Asp Met Tyr Ser Asn
1 5 10 15

Ala Val His Thr Ala Arg Asp Tyr Tyr Asn Ser Glu Asp Ala Ser Asn
20 25 30

Phe Tyr Tyr His Val
35

<210> 10

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide

<400> 10

Gly Ser Val Leu Phe Thr Asp Pro Met Ala Ser Asp Asp Ala Lys
1 5 10 15

<210> 11

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide

<400> 11

Thr Gly Leu Arg Asn Tyr Gln Ala Gly Asn

1 5 10

<210> 12

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide

<400> 12

Leu Xaa Glu Leu Gly Pro Ile Leu Asp Arg Leu His Leu Asp Ser Gly

1 5 10 15

<210> 13

<211> 25

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide

<400> 13

Glu Leu Thr Arg Leu Gly Leu Gln Asn Ile Glu Phe Glu Asp Leu Ser

1 5 10 15

Glu Tyr Leu Pro Val His Tyr Gly Arg
20 25

<210> 14
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: peptide

<400> 14
Val Asp Ile Ser Pro Glu Thr Arg Ile Leu Asp Leu Gly Ser Gly Tyr
1 5 10 15

Gly Ala

<210> 15
<211> 23
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: peptide

<220>
<221> SITE
<222> (4)
<223> Xaa = Thr or Glu

<400> 15

Asn Thr Thr Xaa Glu Gln Asp Phe Gly Ala Asp Pro Thr Lys Val Arg
1 5 10 15

Asp Thr Asp Ala Tyr Thr Glu
20

<210> 16
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: peptide

<400> 16
Val Arg Asp Thr Asp His Tyr Thr Glu Glu Tyr Val Asp
1 5 10

<210> 17
<211> 14
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: peptide

<400> 17
Asp Tyr Thr Arg Arg Leu Met His Glu Val Gly Phe Gln Lys
1 5 10

<210> 18

<211> 17

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide

<400> 18

Ala Thr Tyr Arg Asp Ala Asp Pro Asp Phe Phe Leu His Val Ala Glu
1 5 10 15

Lys

<210> 19

<211> 24

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide

<400> 19

Val Arg Asp Thr Asp His Tyr Thr Glu Glu Tyr Val Asp Gly Phe Val
1 5 10 15

Asp Lys Trp Asp Asp Leu Ile Asp
20

<210> 20

<211> 22

<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<220>
<221> modified_base
<222> (15)
<223> i

<400> 20
gargaygart tygtngayat gt 22

<210> 21
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> ytgrttdatytcraaytcrtc 21

<210> 22
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<220>

<221> modified_base
<222> (15)
<223> i

<400> 22
garcargayt tyggngcnga ycc 23

<210> 23
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<220>
<221> modified_base
<222> (12)
<223> i

<400> 23
arraaraart cnggrtcngc rtc 23

<210> 24
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 24

cggaccatgg atacgactac tgagcag 27

<210> 25
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 25
gctcagatct gtcctcctcc cgatattcct tctc 34

<210> 26
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 26
gcatgccatg gcgacgcgct acgacgatca a 31

<210> 27
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 27

gggaagatct ccctttgcgg aagtaaaaga tacc 34

<210> 28

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 28

gctgccatgg agaagagcta ccgcaccgag 30

<210> 29

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 29

gggaagatct tgccctggcg tggatgatgc ccca 34

<210> 30

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 30

catgccatgg ccaagagcgt ggacgatctt 30